

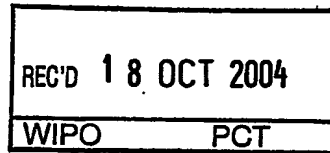


EP04/11082



INVESTOR IN PEOPLE

**PRIORITY
DOCUMENT**
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)



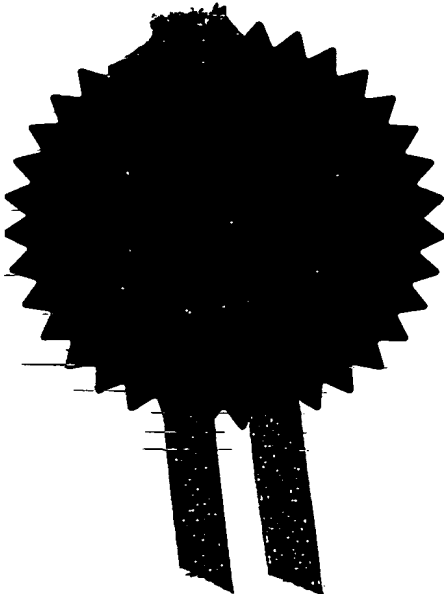
The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.



Signed

Dated 18 August 2004

The
Patent
Office

1/77

030CT03 E841777-1 D02029
P01/7700 0-00-0323112.3

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)



The Patent Office
Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference

CJ/KMM/VB60452

2. Patent application number

(The Patent Office will fill in his part)

0323112.3

- 2 OCT 2003

3. Full name, address and postcode of the or of each applicant (underline all surnames)

GlaxoSmithKline Biologicals s.a.
Rue de l'Institut 89, B-1330 Rixensart, , Belgium

Patents ADP number (*if you know it*)

8101271001

Belgian

If the applicant is a corporate body, give the country/state of its incorporation

4. Title of the invention

Novel Compounds

5. Name of your agent (*if you have one*)

Corporate Intellectual Property

"Address for service" in the United Kingdom to which all correspondence should be sent

(including the postcode)

GlaxoSmithKline
Corporate Intellectual Property (CN9 25.1)
980 Great West Road
BRENTFORD
Middlesex TW8 9GS

Patents ADP number (*if you know it*)

7960982004

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or each of these earlier applications and (*if you know it*) the or each application number

Country	Priority application number (<i>if you know it</i>)	Date of filing (<i>day / month / year</i>)
---------	--	---

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application

Date of filing
(*day / month / year*)

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? (*Answer yes if:*

- a) any applicant named in part 3 is not an inventor, or
 - b) there is an inventor who is named as an applicant, or
 - c) any named applicant is a corporate body
- See note (d)

Enter the number of sheets for any of the following items you are filing with this form.
Do not count copies of the same document

Continuation sheets of this form	-
Description	138
Claim(s)	10
Abstract	1
Drawings	3

only *ll*

10. If you are also filing any of the following, state how many against each item.

Priority Documents

Translations of priority documents

Statement of inventorship and right to grant of a patent (*Patents Form 7/77*)

Request for preliminary examination and search (*Patents Form 9/77*)

Request for substantive examination (*Patents Form 10/77*)

Any other documents
(*please specify*)

11.

We request the grant of a patent on the basis of this application

Signature *Michael Lubinski* Date 2-Oct-03
M J Lubinski

12. Name and daytime telephone number of person to contact in the United Kingdom

M J Lubinski 020 80474434

Warning

After an application for a Patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission unless an application has been filed at least six weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

- a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505
- b) Write your answers in capital letters using black ink or you may type them.
- c) If there is not enough space for all relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.
- d) If you have answered 'Yes' Patents Form 7/77 will need to be filed.
- f) For details of the fee and ways to pay please contact the Patent Office.

Novel Compounds

FIELD OF THE INVENTION

This invention relates to polynucleotides, (herein referred to as "BASB232 polynucleotide(s)"), polypeptides encoded by them (referred to herein as "BASB232" or "BASB232 polypeptide(s)"), recombinant materials and methods for their production. In particular, the invention relates to immunogenic compositions and vaccines containing single polypeptide or nucleotides or advantageous combinations thereof. In another aspect of the invention, the invention relates to methods for using such polypeptides and polynucleotides for the treatment or prevention of *Bordetella* infections. In a further aspect, the invention relates to diagnostic assays for detecting *Bordetella* infection.

BACKGROUND OF THE INVENTION

The bacterium *Bordetella pertussis* is the causative agent for whooping cough, a respiratory disease that can be severe in infants and young children. The clinical course of the disease is characterised by paroxysms of rapid coughs followed by inspiratory effort, often associated with a characteristic 'whooping' sound. In serious cases, oxygen deprivation can lead to brain damage, however the most common complication is secondary pneumonia.

20

Whooping cough is usually considered to be caused by *B. pertussis*, but occasionally *B. parapertussis* is isolated from patients with typical signs and symptoms of whooping cough. *B. parapertussis* infection is of lower frequency than *B. pertussis* with 5-10% of whooping cough being associated with *B. parapertussis* (Mertsola (1985) Eur J Clin Microbiol 4; 123; Lautrop (1971) Lancet 1(7711) 1195-1198). *B. parapertussis* is associated with mild clinical symptoms which, combined with its serological cross-reactivity with *B. pertussis*, makes *B. parapertussis* difficult to diagnose.

25

The first generation of vaccines against *B. pertussis* were whole cell vaccines, composed of whole killed bacteria. These were introduced in many countries in the

30

1950s and 1960s and were successful at reducing the incidence of whooping cough. A problem with whole cell *B. pertussis* vaccines is the high level of reactogenicity associated with them. Acellular vaccines containing purified *B. pertussis* proteins are less reactogenic and have been adopted for the vaccination programmes of many countries. Acellular vaccines typically containing pertussis toxin (PT), filamentous haemagglutinin (FHA) and quite often pertactin (PRN), are widely used and provide effective protection from the severity of whooping cough.

Despite vaccination, whooping cough remains an endemic disease (Mooi et al (2001) Emerging Infectious Diseases 7; 526). Whooping cough has re-emerged in Australia, Canada and The Netherlands; countries with highly vaccinated populations. A comparison of pre-vaccination strains with strains isolated recently, has shown antigenic drift, particularly in PT and PRN (Mooi et al (1998) Infection and Immunity 66; 670). It is widely acknowledged that current vaccines protect against severe disease but do not eliminate *Bordetella pertussis* from the body (Cherry et al (1998) Vaccine 16; 1901, Hewlett and Halperin (1998) Vaccine 16; 1899, Storsaeter et al (1998) Vaccine 16; 1907). The defence mechanisms of *Bordetella pertussis* allow it to evade elimination from the body, indicating that current vaccines do not completely disable these defence mechanisms.

Vaccination using whole cell *B. pertussis* vaccines (Pw), appears to protect against *B. parapertussis* infection, probably due to the similarity of the two bacteria. *B. parapertussis* infection in unvaccinated infants may lead to severe and fatal complications, whereas in individuals vaccinated with Pw, a milder, often subclinical course of whooping cough is seen (Long et al (1990) Pediatric Infect Dis J 9; 700). Theoretically, the introduction of acellular pertussis vaccines containing only two or three purified proteins could reduce the ability of vaccination to protect against *B. parapertussis*.

Accordingly, further improved acellular vaccines against whooping cough are required that combine low reactogenicity with an ability to elicit a protective response against *Bordetella*, particularly both *B. pertussis* and *B. parapertussis*, infection. The identification of new candidate antigens and particularly effective combinations of
5 antigens will allow the development of such vaccines.

SUMMARY OF THE INVENTION

The present invention relates to immunogenic compositions containing BASB232, in particular BASB232 polypeptides and BASB232 polynucleotides, recombinant materials
10 and methods for their production. In a further aspect, the invention relates to combination of polypeptides and nucleotides that interact advantageously in the prevention or treatment of microbial disease. In another aspect, the invention relates to methods for using such polypeptides, polynucleotides and combinations, including prevention and treatment of *Bordetella* diseases, amongst others. In a further aspect, the invention relates to diagnostic
15 assays for detecting diseases associated with microbial infections and conditions associated with such infections, such as assays for detecting expression or activity of BASB232 polynucleotides or polypeptides.

Various changes and modifications within the spirit and scope of the disclosed invention
20 will become readily apparent to those skilled in the art from reading the following descriptions and from reading the other parts of the present disclosure.

DESCRIPTION OF FIGURES

Figure 1 – is a graph showing protection against challenge with *B. pertussis* strain
25 Tohama in groups of mice pre-immunised with carrier DT BrkA, DTPa-2, DTPa-2 BrkA, DTPa-3 or DTPa-3 BrkA. Results are expressed as the number of CFU isolated per lung at different time points after challenge.

Figure 2 – is a graph showing protection against challenge with *B. pertussis* strain 18323 in groups of mice pre-immunised with carrier DT BrkA, DTPa-2, DTPa-2 BrkA, DTPa-3 or DTPa-3 BrkA. Results are expressed as the number of CFU isolated per lung at different time points after challenge.

5

Figure 3 – graphs showing protection against challenge with *B. pertussis* or *B. parapertussis* in groups of mice preimmunised with DTPw or DTPa from several sources. Results are expressed as number of CFU isolated from the lung at different time points after challenge.

10

DESCRIPTION OF THE INVENTION

The invention relates to BASB232 polypeptides and polynucleotides as described in greater detail below. In particular, the invention relates to polypeptides and polynucleotides of BASB232 of *B. pertussis*, particularly comprised in immunogenic compositions.

15

The invention relates especially to BASB232 polynucleotides and encoded polypeptides listed in table 1. Those polynucleotides and encoded polypeptides have the nucleotide and amino acid sequences set out in SEQ ID NO:1 to SEQ ID NO:98 as described in table 1.

Table 1

Name	Length (nT)	Length (aa)	SEQ ID nucl.	SEQ ID prot.	Description
Orf1	2211	737	1	2	Ferric enterobactin receptor (BfeA), <i>Bordetella pertussis</i> (95%)
Orf2	2475	812	3	4	Probable hydroxamate-type ferrisiderophore receptor (BfrB), <i>Pseudomonas aeruginosa</i> (40%)
Orf3	2403	729	5	6	Putative hydroxamate-type ferrisiderophore receptor signal peptide protein (BfrC), <i>Pseudomonas solanacearum</i> (38%)
Orf4	2304	734	7	8	Putative ferric siderophore receptor (FauA), <i>Bordetella bronchiseptica</i> (97%)
Orf5	2187	825	9	10	Unidentified ferric siderophore receptor, <i>Bordetella bronchiseptica</i> (94%)

Orf6	2064	801	11	12	Ferric alcaligin siderophore receptor, <i>Bordetella pertussis</i> (100%)
Orf7	2229	743	13	14	Hydroxamate-type ferrisidero-phore receptor (iron transport protein fiu), <i>Pseudomonas aeruginosa</i> (37%)
Orf8	2268	756	15	16	Hydroxamate-type ferrisidero-phore receptor (iron transport protein fiu), <i>Pseudomonas aeruginosa</i> (41%)
Orf9	2106	702	17	18	Putative hydroxamate-type ferrisiderophore receptor signal peptide protein, <i>Pseudomonas solanacearum</i> (40%)
Orf10	2610	870	19	20	BhuR, outer membrane heme receptor, <i>Bordetella pertussis</i> (100%)
Orf11	2280	760	21	22	Probable tonb-dependent receptor, <i>Pseudomonas aeruginosa</i> (34%)
Orf12	1887	629	23	24	Probable tonb-dependent receptor, <i>Pseudomonas aeruginosa</i> (34%)
Orf13	1731	577	25	26	Ferrisiderophore receptor-like protein, <i>Pseudomonas sp</i> (57%)
Orf14	1434	478	27	28	Probable tonB-dependant receptor Yncd precursor, <i>Escherichia coli</i> (56%)
Orf15	2730	910	29	30	Pertactin outer membrane protein, <i>Bordetella pertussis</i> (100%)
Orf16	2748	915	31	32	Vag8 protein, <i>Bordetella pertussis</i> (96%)
Orf17	3033	1010	33	34	BrkA, <i>Bordetella pertussis</i> (81%)
Orf18	1944	647	35	36	Tcf protein, <i>Bordetella pertussis</i> (74%)
Orf19	1245	418	37	38	Phg protein, <i>Bordetella pertussis</i> (81%)
Orf20	2712	903	39	40	BapA protein, <i>Bordetella pertussis</i> (85%)
Orf21	1446	482	41	42	BapB protein, <i>Bordetella pertussis</i> (87%)
Orf22	2277	759	43	44	Putative autotransporter BapC, <i>Bordetella pertussis</i> (86%)
Orf23	1545	515	45	46	Pertactin-like protein, <i>Bordetella pertussis</i> (47%)
Orf24	1191	397	47	48	Tcf-like protein, <i>Bordetella pertussis</i> (56%)
Orf25	6903	2300	49	50	Extracellular serine protease, <i>Brucella melitensis</i> (25%)
Orf26	2622	873	51	52	Autotransporter protein, <i>Agrobacterium tumefaciens</i> (43%)
Orf27	3120	1039	53	54	Autotransporter subtilisin-like protease (SphB1), <i>Bordetella pertussis</i> (93%)
Orf28	2241	747	55	56	Heme/hemopexin utilization protein c precursor, <i>Haemophilus influenzae</i> (48%)
Orf29	1575	525	57	58	Lipoprotein (piln protein), <i>Escherichia coli</i> (22%)

Orf30	1509	503	59	60	Immunogenic protein, <i>Deinococcus radiodurans</i> (35%)
Orf31	1491	497	61	62	Probable outer membrane lipoprotein precursor, <i>Pseudomonas aeruginosa</i> (48%)
Orf32	1491	497	63	64	Probable outer membrane efflux protein precursor, <i>Pseudomonas aeruginosa</i> (43%)
Orf33	1380	460	65	66	Oprm, <i>Pseudomonas aeruginosa</i> (45%)
Orf34	1347	449	67	68	Probable outer membrane channel signal peptide protein, <i>Ralstonia solanacearum</i> (40%)
Orf35	1287	429	69	70	Putative membrane-bound lytic murein transglycosylase a transmembrane protein (MltA), <i>Ralstonia solanacearum</i> (42%)
Orf36	1143	381	71	72	Putative membrane-bound lytic murein transglycosylase b protein (MltB), <i>Ralstonia solanacearum</i> (40%)
Orf37	1095	365	73	74	Putative polysaccharide export protein yccz precursor, <i>Escherichia coli</i> (34%)
Orf38	897	299	75	76	Putative serine protease transmembrane protein, <i>Ralstonia solanacearum</i> (55%)
Orf39	852	284	77	78	Hypothetical protein pa4632, <i>Pseudomonas aeruginosa</i> (52%)
Orf40	846	282	79	80	Competence lipoprotein coml precursor, <i>Neisseria meningitidis</i> (45%)
Orf41	813	271	81	82	Probable lipoprotein precursor (vacj) transmembrane, <i>Ralstonia solanacearum</i> (43%)
Orf42	801	267	83	84	Putative outer membrane lipoprotein, <i>Salmonella typhimurium</i> (24%)
Orf43	690	230	85	86	Flagellar l-ring protein precursor (basal body l-ring protein), <i>Escherichia coli</i> (51%)
Orf44	678	226	87	88	Hypothetical lipoprotein ydcl precursor, <i>Escherichia coli</i> (32%)
Orf45	558	186	89	90	Probable peptidoglycan-associated lipoprotein precursor (Pal), <i>Ralstonia solanacearum</i> (63%)
Orf46	552	184	91	92	Putative outer membrane lipoprotein (OmlA), <i>Bordetella pertussis</i> (100%)
Orf47	546	182	93	94	Hypothetical transmembrane protein smc00354, <i>Rhizobium meliloti</i> (36%)
Orf48	501	167	95	96	Putative outer membrane lipoprotein transmembrane, <i>Ralstonia solanacearum</i> (40%)
Orf49	456	152	97	98	Lipoprotein, <i>Vibrio cholerae</i> (44%)

The percentage shown in table 1 are the identity percentage shared by each sequence of the BASB232 polypeptides and their homologous polypeptides found in *B.pertussis* or in other organisms (by a BLAST homology search).

- 5 It is understood that sequences recited in the Sequence Listing below as “DNA” represent an exemplification of one embodiment of the invention, since those of ordinary skill will recognize that such sequences can be usefully employed in polynucleotides in general, including ribopolynucleotides.

10 The sequences of the BASB232 polynucleotides are set out in SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97. SEQ Group 1 refers herein to any one of the polynucleotides set out in SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97.

15 The sequences of the BASB232 encoded polypeptides are set out in SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98. SEQ Group 2 refers herein to any one of the encoded polypeptides set out in SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98.

The BASB232 polynucleotides set out in SEQ ID 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25 and 27 belong to the iron transporter protein family.

25 The BASB232 polynucleotides set out in SEQ ID 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 and 53 belong to the autotransporter proteins family.

The BASB232 polynucleotides set out in SEQ ID 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95 and 97 belong to the lipoproteins family.

30 The BASB232 polypeptides set out in SEQ ID 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26 and 28 belong to the iron transporter protein family.

The BASB232 polypeptides set out in SEQ ID 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 and 54 belong to the autotransporter proteins family.

The BASB232 polypeptides set out in SEQ ID 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96 and 98 belong to the lipoproteins family.

5

Polypeptides

In one aspect of the invention there are provided polypeptides of *B. pertussis* referred to herein as "BASB232" and "BASB232 polypeptides" as well as biologically, diagnostically, prophylactically, clinically or therapeutically useful variants thereof, and compositions, preferably immunogenic compositions comprising the same.

10

The present invention further provides for:

(a) an isolated polypeptide which comprises an amino acid sequence which has at least 85% identity, preferably at least 90% identity, more preferably at least 95% identity, most preferably at least 97-99% or exact identity, to that of any sequence of SEQ Group 2;

15

(b) a polypeptide encoded by an isolated polynucleotide comprising a polynucleotide sequence which has at least 85% identity, preferably at least 90% identity, more preferably at least 95% identity, even more preferably at least 97-99% or exact identity to any sequence of SEQ Group 1 over the entire length of the selected sequence of SEQ Group 1;

20

or

(c) a polypeptide encoded by an isolated polynucleotide comprising a polynucleotide sequence encoding a polypeptide which has at least 85% identity, preferably at least 90% identity, more preferably at least 95% identity, even more preferably at least 97-99% or exact identity, to the amino acid sequence of any sequence of SEQ Group 2.

25

The BASB232 polypeptides provided in SEQ Group 2 are the BASB232 polypeptides from *B. pertussis* as described in table 1.

The invention also provides an immunogenic fragment of a BASB232 polypeptides, that is, a contiguous portion of the BASB232 polypeptide which has the same or substantially

30

the same immunogenic activity as the polypeptide comprising the corresponding amino acid sequence selected from SEQ Group 2 ; That is to say, the fragment (if necessary when coupled to a carrier) is capable of raising an immune response which recognises the BASB232 polypeptide. Such an immunogenic fragment may include, for example, the

5 BASB232 polypeptide lacking an N-terminal leader sequence, and/or a transmembrane domain and/or a C-terminal anchor domain. In a preferred aspect the immunogenic fragment of BASB232 according to the invention comprises substantially all of the extracellular domain of a polypeptide which has at least 85% identity, preferably at least 90% identity, more preferably at least 95% identity, most preferably at least 97-99%

10 identity, to that a sequence selected from SEQ Group 2 over the entire length of said sequence.

A fragment is a polypeptide having an amino acid sequence that is entirely the same as part but not all of any amino acid sequence of any polypeptide of the invention. As with

15 BASB232 polypeptides, fragments may be "free-standing," or comprised within a larger polypeptide of which they form a part or region, most preferably as a single continuous region in a single larger polypeptide.

Preferred fragments include, for example, truncation polypeptides having a portion of an amino acid sequence selected from SEQ Group 2 or of variants thereof, such as a continuous series of residues that includes an amino- and/or carboxyl-terminal amino acid sequence. Degradation forms of the polypeptides of the invention produced by or in a host cell, are also preferred. Further preferred are fragments characterized by structural or functional attributes such as fragments that comprise alpha-helix and alpha-helix forming regions,

20 beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions.

25

Further preferred fragments include an isolated polypeptide comprising an amino acid sequence having at least 15, 20, 30, 40, 50 or 100 contiguous amino acids from the amino acid sequence selected from SEQ Group 2 or an isolated polypeptide comprising an amino acid sequence having at least 15, 20, 30, 40, 50 or 100 contiguous amino acids truncated or deleted from the amino acid sequence selected from SEQ Group 2 .

The BASB232 polypeptides set out in SEQ ID 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 53 and 54 belong to an autotransporter proteins family. In this family, there are two domains : the passenger domain that is surface exposed and the beta domain that is anchored in the outer membrane protein. The passenger domain is a preferred fragment for vaccine use. The passenger domain was predicted for each BASB232 polypeptides set out in SEQ ID 30, 32, 34, 36, 38, 40, 42, 44, 50, and 52 in table 2.

Table 2

encoded peptidic sequence	1 st amino acids of the preferred fragment	Last amino acids of the preferred fragment
SEQ ID NO:30	35	604
SEQ ID NO:32	40	614
SEQ ID NO:34	41	706
SEQ ID NO:36	40	132
SEQ ID NO:38	36	114
SEQ ID NO:40	31	595
SEQ ID NO:42	1	185
SEQ ID NO:44	1	458
SEQ ID NO:50	38	1984
SEQ ID NO:52	43	561

Fragments described in table 2 are preferred fragments. These fragments may be readily modified by adding or removing 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40 or 50 amino acids from either or both N and C termini.

Still further preferred fragments are those which comprise a B-cell or T-helper epitope, for example those fragments/peptides described in Example 8.

- 5 Fragments of the polypeptides of the invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, these fragments may be employed as intermediates for producing the full-length polypeptides of the invention.

10 Particularly preferred are variants in which several, 5-10, 1-5, 1-3, 1-2 or 1 amino acids are substituted, deleted, or added in any combination.

The polypeptides, or immunogenic fragments, of the invention may be in the form of the "mature" protein or may be a part of a larger protein such as a precursor or a fusion protein. It is often advantageous to include an additional amino acid sequence which
15 contains secretory or leader sequences, pro-sequences, sequences which aid in purification such as multiple histidine residues, or an additional sequence for stability during recombinant production. Furthermore, addition of exogenous polypeptide or lipid tail or polynucleotide sequences to increase the immunogenic potential of the final molecule is also considered.

20

In one aspect, the invention relates to genetically engineered soluble fusion proteins comprising a polypeptide of the present invention, or a fragment thereof, and various portions of the constant regions of heavy or light chains of immunoglobulins of various subclasses (IgG, IgM, IgA, IgE). Preferred as an immunoglobulin is the constant part of
25 the heavy chain of human IgG, particularly IgG1, where fusion takes place at the hinge region. In a particular embodiment, the Fc part can be removed simply by incorporation of a cleavage sequence which can be cleaved with blood clotting factor Xa.

30 Furthermore, this invention relates to processes for the preparation of these fusion proteins by genetic engineering, and to the use thereof for drug screening, diagnosis and

therapy. A further aspect of the invention also relates to polynucleotides encoding such fusion proteins. Examples of fusion protein technology can be found in International Patent Application Nos. WO94/29458 and WO94/22914.

- 5 The proteins may be chemically conjugated, or expressed as recombinant fusion proteins allowing increased levels to be produced in an expression system as compared to non-fused protein. The fusion partner may assist in providing T helper epitopes (immunological fusion partner), preferably T helper epitopes recognised by humans, or assist in expressing the protein (expression enhancer) at higher yields than the native
10 recombinant protein. Preferably the fusion partner will be both an immunological fusion partner and expression enhancing partner.

Fusion partners include protein D from *Haemophilus influenzae* and the non-structural protein from influenza virus, NS1 (hemagglutinin). Another fusion partner is the protein
15 known as Omp26 (WO 97/01638). Another fusion partner is the protein known as LytA. Preferably the C terminal portion of the molecule is used. LytA is derived from *Streptococcus pneumoniae* which synthesize an N-acetyl-L-alanine amidase, amidase LytA, (coded by the *lytA* gene {Gene, 43 (1986) page 265-272}) an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal
20 domain of the LytA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of *E.coli* C-LytA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LytA fragment at its amino terminus has been described {Biotechnology: 10, (1992) page 795-798}. It is possible to use the
25 repeat portion of the LytA molecule found in the C terminal end starting at residue 178, for example residues 188 - 305.

The present invention also includes variants of the aforementioned polypeptides, that is polypeptides that vary from the referents by conservative amino acid substitutions,
30 whereby a residue is substituted by another with like characteristics. Typical such

substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; and among the basic residues Lys and Arg; or aromatic residues Phe and Tyr.

5 Polypeptides of the present invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

10

It is most preferred that a polypeptide of the invention is derived from *B. pertussis*, however, it may preferably be obtained from other organisms of the same taxonomic genus. A polypeptide of the invention may also be obtained, for example, from organisms of the same taxonomic family or order.

15

Polynucleotides

It is an object of the invention to provide polynucleotides that encode BASB232 polypeptides, particularly polynucleotides that encode polypeptides herein designated BASB232.

20

In a particularly preferred embodiment of the invention the polynucleotide comprises a region encoding BASB232 polypeptides comprising sequences set out in SEQ Group 1 which include full length gene, or a variant thereof.

25 Polynucleotides of the invention do not encompass a complete genomic DNA from a *Bordetella* species, preferably *B. pertussis* or *B. parapertussis*.

As a further aspect of the invention there are provided isolated nucleic acid molecules encoding and/or expressing BASB232 polypeptides and polynucleotides, particularly *B.*

30 *pertussis* BASB232 polypeptides and polynucleotides, including, for example,

unprocessed RNAs, ribozyme RNAs, mRNAs, cDNAs, B- and Z-DNAs. Further embodiments of the invention include biologically, diagnostically, prophylactically, clinically or therapeutically useful polynucleotides and polypeptides, and variants thereof, and compositions, preferably immunogenic compositions, comprising the same.

5

Another aspect of the invention relates to isolated polynucleotides, including at least one full length gene, that encode BASB232 polypeptides having a deduced amino acid sequence of SEQ Group 2 and polynucleotides closely related thereto and variants thereof.

- 10 In another particularly preferred embodiment of the invention relates to BASB232 polypeptides from *B. pertussis* comprising or consisting of an amino acid sequence selected from SEQ Group 2 or a variant thereof.

- Using the information provided herein, such as a polynucleotide sequence set out in SEQ Group 1, a polynucleotide of the invention encoding BASB232 polypeptide may be obtained using standard cloning and screening methods, such as those for cloning and sequencing chromosomal DNA fragments from bacteria using *B. pertussis* strain Tohama I cells as starting material, followed by obtaining a full length clone. For example, to obtain a polynucleotide sequence of the invention, such as a polynucleotide sequence given in SEQ Group 1, typically a library of clones of chromosomal DNA of *B. pertussis* strain Tohama I in *E. coli* or some other suitable host is probed with a radiolabeled oligonucleotide, preferably a 17-mer or longer, derived from a partial sequence. Clones carrying DNA identical to that of the probe can then be distinguished using stringent hybridization conditions. By sequencing the individual clones thus identified by hybridization with sequencing primers designed from the original polypeptide or polynucleotide sequence it is then possible to extend the polynucleotide sequence in both directions to determine a full length gene sequence. Conveniently, such sequencing is performed, for example, using denatured double stranded DNA prepared from a plasmid clone. Suitable techniques are described by Maniatis, T., Fritsch, E.F. and Sambrook et al., *MOLECULAR CLONING, A LABORATORY MANUAL*, 2nd Ed.; Cold Spring Harbor
- 15
- 20
- 25
- 30

Laboratory Press, Cold Spring Harbor, New York (1989). (see in particular Screening By Hybridization 1.90 and Sequencing Denatured Double-Stranded DNA Templates 13.70).

Direct genomic DNA sequencing may also be performed to obtain a full length gene sequence. Illustrative of the invention, each polynucleotide set out in SEQ Group 1 was
5 discovered in a DNA library derived from *B. pertussis*.

Moreover, each DNA sequence set out in SEQ Group 1 contains an open reading frame encoding a protein having about the number of amino acid residues set forth in SEQ Group
2 with a deduced molecular weight that can be calculated using amino acid residue
10 molecular weight values well known to those skilled in the art.

The polynucleotides of SEQ Group 1, between the start codon and the stop codon, encode respectively the polypeptides of SEQ Group 2. The nucleotide number of start codon and first nucleotide of the stop codon are listed in table 3 for each polynucleotide of SEQ Group

15 1.

Table 3

Name	Start codon	1 st nucleotide of stop codon
Orf1	1	2212
Orf2	1	2476
Orf3	1	2404
Orf4	1	2305
Orf5	1	2188
Orf6	1	2065
Orf7	1	2230
Orf8	1	2269
Orf9	1	2107
Orf10	1	2611
Orf11	1	2281
Orf12	1	1888
Orf13	1	1732
Orf14	1	1435
Orf15	1	2731

Orf16	1	2746
Orf17	1	3031
Orf18	1	1942
Orf19	1	1255
Orf20	1	2710
Orf21	1	1447
Orf22	1	2278
Orf23	1	1546
Orf24	1	1192
Orf25	1	6901
Orf26	1	2620
Orf27	1	3118
Orf28	1	2242
Orf29	1	1576
Orf30	1	1510
Orf31	1	1492
Orf32	1	1492
Orf33	1	1381
Orf34	1	1348
Orf35	1	1288
Orf36	1	1144
Orf37	1	1096
Orf38	1	898
Orf39	1	853
Orf40	1	847
Orf41	1	814
Orf42	1	802
Orf43	1	691
Orf44	1	679
Orf45	1	559
Orf46	1	553
Orf47	1	547
Orf48	1	502
Orf49	1	457

In a further aspect, the present invention provides for an isolated polynucleotide comprising or consisting of:

- (a) a polynucleotide sequence which has at least 85% identity, preferably at least 90% identity, more preferably at least 95% identity, even more preferably at least 97-99% or exact identity to any sequence from SEQ Group 1 over the entire length of the polynucleotide sequence from SEQ Group 1; or
- 5 (b) a polynucleotide sequence encoding a polypeptide which has at least 85% identity, preferably at least 90% identity, more preferably at least 95% identity, even more preferably at least 97-99% or 100% exact, to any amino acid sequence selected from SEQ Group 2, over the entire length of the amino acid sequence from SEQ Group 2.
- 10 A polynucleotide encoding a polypeptide of the present invention, including homologs and orthologs from species other than *B. pertussis*, may be obtained by a process which comprises the steps of screening an appropriate library under stringent hybridization conditions (for example, using a temperature in the range of 45 – 65°C and an SDS concentration from 0.1 – 1%) with a labeled or detectable probe consisting of or comprising
- 15 any sequence selected from SEQ Group 1 or a fragment thereof; and isolating a full-length gene and/or genomic clones containing said polynucleotide sequence.

The invention provides a polynucleotide sequence identical over its entire length to a coding sequence (open reading frame) set out in SEQ Group 1. Also provided by the invention is a

20 coding sequence for a mature polypeptide or a fragment thereof, by itself as well as a coding sequence for a mature polypeptide or a fragment in reading frame with another coding sequence, such as a sequence encoding a leader or secretory sequence, a pre-, or pro- or prepro-protein sequence. The polynucleotide of the invention may also contain at least one non-coding sequence, including for example, but not limited to at least one non-coding 5'

25 and 3' sequence, such as the transcribed but non-translated sequences, termination signals (such as rho-dependent and rho-independent termination signals), ribosome binding sites, Kozak sequences, sequences that stabilize mRNA, introns, and polyadenylation signals. The polynucleotide sequence may also comprise additional coding sequence encoding additional amino acids. For example, a marker sequence that facilitates purification of the

30 fused polypeptide can be encoded. In certain embodiments of the invention, the marker

sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, *Proc. Natl. Acad. Sci., USA* 86: 821-824 (1989), or an HA peptide tag (Wilson *et al.*, *Cell* 37: 767 (1984), both of which may be useful in purifying polypeptide sequence fused to them. Polynucleotides of the invention also include, but are not limited to, polynucleotides comprising a structural gene and its naturally associated sequences that control gene expression.

The nucleotide sequences encoding the BASB232 polypeptides of SEQ Group 2 may be identical to the corresponding polynucleotide encoding sequences of SEQ Group 1. The position of the first and last nucleotides of the encoding sequences of SEQ Group 1 are listed in table 4. Alternatively it may be any sequence, which as a result of the redundancy (degeneracy) of the genetic code, also encodes polypeptides of SEQ Group 2.

Table 4

nucleotidic sequence	encoded peptidic sequence	Start codon	Last nucleotide of encoding sequence
SEQ ID NO:1	SEQ ID NO:2	1	2211
SEQ ID NO:3	SEQ ID NO:4	1	2475
SEQ ID NO:5	SEQ ID NO:6	1	2403
SEQ ID NO:7	SEQ ID NO:8	1	2304
SEQ ID NO:9	SEQ ID NO:10	1	2187
SEQ ID NO:11	SEQ ID NO:12	1	2064
SEQ ID NO:13	SEQ ID NO:14	1	2229
SEQ ID NO:15	SEQ ID NO:16	1	2268
SEQ ID NO:17	SEQ ID NO:18	1	2106
SEQ ID NO:19	SEQ ID NO:20	1	2610
SEQ ID NO:21	SEQ ID NO:22	1	2280
SEQ ID NO:23	SEQ ID NO:24	1	1887
SEQ ID NO:25	SEQ ID NO:26	1	1731

SEQ ID NO:27	SEQ ID NO:28	1	1434
SEQ ID NO:29	SEQ ID NO:30	1	2730
SEQ ID NO:31	SEQ ID NO:32	1	2745
SEQ ID NO:33	SEQ ID NO:34	1	3030
SEQ ID NO:35	SEQ ID NO:36	1	1941
SEQ ID NO:37	SEQ ID NO:38	1	1254
SEQ ID NO:39	SEQ ID NO:40	1	2709
SEQ ID NO:41	SEQ ID NO:42	1	1446
SEQ ID NO:43	SEQ ID NO:44	1	2277
SEQ ID NO:45	SEQ ID NO:46	1	1545
SEQ ID NO:47	SEQ ID NO:48	1	1191
SEQ ID NO:49	SEQ ID NO:50	1	6900
SEQ ID NO:51	SEQ ID NO:52	1	2619
SEQ ID NO:53	SEQ ID NO:54	1	3117
SEQ ID NO:55	SEQ ID NO:56	1	2241
SEQ ID NO:57	SEQ ID NO:58	1	1575
SEQ ID NO:59	SEQ ID NO:60	1	1509
SEQ ID NO:61	SEQ ID NO:62	1	1491
SEQ ID NO:63	SEQ ID NO:64	1	1491
SEQ ID NO:65	SEQ ID NO:66	1	1380
SEQ ID NO:67	SEQ ID NO:68	1	1347
SEQ ID NO:69	SEQ ID NO:70	1	1287
SEQ ID NO:71	SEQ ID NO:72	1	1143
SEQ ID NO:73	SEQ ID NO:74	1	1095
SEQ ID NO:75	SEQ ID NO:76	1	897
SEQ ID NO:77	SEQ ID NO:78	1	852
SEQ ID NO:79	SEQ ID NO:80	1	846
SEQ ID NO:81	SEQ ID NO:82	1	813

SEQ ID NO:83	SEQ ID NO:84	1	801
SEQ ID NO:85	SEQ ID NO:86	1	690
SEQ ID NO:87	SEQ ID NO:88	1	678
SEQ ID NO:89	SEQ ID NO:90	1	558
SEQ ID NO:91	SEQ ID NO:92	1	552
SEQ ID NO:93	SEQ ID NO:94	1	546
SEQ ID NO:95	SEQ ID NO:96	1	501
SEQ ID NO:97	SEQ ID NO:98	1	456

The term "polynucleotide encoding a polypeptide" as used herein encompasses polynucleotides that include a sequence encoding a polypeptide of the invention, particularly a bacterial polypeptide and more particularly a polypeptide of the *B. pertussis* BASB232

5 having an amino acid sequence set out in any of the sequences of SEQ Group 2. The term also encompasses polynucleotides that include a single continuous region or discontinuous regions encoding the polypeptide (for example, polynucleotides interrupted by integrated phage, an integrated insertion sequence, an integrated vector sequence, an integrated transposon sequence, or due to RNA editing or genomic DNA reorganization) together with
10 additional regions, that also may contain coding and/or non-coding sequences.

The invention further relates to variants of the polynucleotides described herein that encode variants of a polypeptides having a deduced amino acid sequence of any of the sequences of SEQ Group 2. Fragments of polynucleotides of the invention may be used, for example, to
15 synthesize full-length polynucleotides of the invention.

Preferred fragments are those polynucleotides which encode a B-cell or T-helper epitope, for example the fragments/peptides described in Example 8, and recombinant, chimeric genes comprising said polynucleotide fragments.

Further particularly preferred embodiments are polynucleotides encoding BASB232 variants, that have the amino acid sequence of BASB232 polypeptides of any sequence from SEQ Group 2 in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, modified, deleted and/or added, in any combination. Especially preferred among these are silent substitutions, additions and deletions, that do not alter the properties and activities of BASB232 polypeptides.

Further preferred embodiments of the invention are polynucleotides that are at least 85% identical over their entire length to polynucleotides encoding BASB232 polypeptides having an amino acid sequence set out in any of the sequences of SEQ Group 2, and polynucleotides that are complementary to such polynucleotides. Alternatively, most highly preferred are polynucleotides that comprise a region that is at least 90% identical over its entire length to polynucleotides encoding BASB232 polypeptides and polynucleotides complementary thereto. In this regard, polynucleotides at least 95% identical over their entire length to the same are particularly preferred. Furthermore, those with at least 97% are highly preferred among those with at least 95%, and among these those with at least 98% and at least 99% are particularly highly preferred, with at least 99% being the more preferred.

Preferred embodiments are polynucleotides encoding polypeptides that retain substantially the same biological function or activity as mature polypeptides encoded by a DNA sequences selected from SEQ Group 1.

In accordance with certain preferred embodiments of this invention there are provided polynucleotides that hybridize, particularly under stringent conditions, to BASB232 polynucleotide sequences, such as those polynucleotides in SEQ Group 1.

The invention further relates to polynucleotides that hybridize to the polynucleotide sequences provided herein. In this regard, the invention especially relates to polynucleotides that hybridize under stringent conditions to the polynucleotides described herein. As herein

used, the terms "stringent conditions" and "stringent hybridization conditions" mean hybridization occurring only if there is at least 95% and preferably at least 97% identity between the sequences. A specific example of stringent hybridization conditions is overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 micrograms/ml of denatured, sheared salmon sperm DNA, followed by washing the hybridization support in 0.1x SSC at about 65°C. Hybridization and wash conditions are well known and exemplified in Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), particularly Chapter 11 therein. Solution hybridization may also be used with the polynucleotide sequences provided by the invention.

The invention also provides a polynucleotide consisting of or comprising a polynucleotide sequence obtained by screening an appropriate library containing the complete gene for a polynucleotide sequence set forth in any of the sequences of SEQ Group 1 under stringent hybridization conditions with a probe having the sequence of said polynucleotide sequence set forth in the corresponding sequences of SEQ Group 1 or a fragment thereof; and isolating said polynucleotide sequence. Fragments useful for obtaining such a polynucleotide include, for example, probes and primers fully described elsewhere herein.

As discussed elsewhere herein regarding polynucleotide assays of the invention, for instance, the polynucleotides of the invention, may be used as a hybridization probe for RNA, cDNA and genomic DNA to isolate full-length cDNAs and genomic clones encoding BASB232 and to isolate cDNA and genomic clones of other genes that have a high identity, particularly high sequence identity, to the BASB232 genes. Such probes generally will comprise at least 15 nucleotide residues or base pairs. Preferably, such probes will have at least 30 nucleotide residues or base pairs and may have at least 50 nucleotide residues or base pairs. Particularly preferred probes will have at least 20 nucleotide residues or base pairs and will have less than 30 nucleotide residues or base pairs.

A coding region of BASB232 genes may be isolated by screening using a DNA sequences provided in SEQ Group 1 to synthesize an oligonucleotide probe. A labeled oligonucleotide having a sequence complementary to that of a gene of the invention is then used to screen a library of cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

There are several methods available and well known to those skilled in the art to obtain full-length DNAs, or extend short DNAs, for example those based on the method of Rapid Amplification of cDNA ends (RACE) (see, for example, Frohman, *et al.*, *PNAS USA* 85: 8998-9002, 1988). Recent modifications of the technique, exemplified by the Marathon™ technology (Clontech Laboratories Inc.) for example, have significantly simplified the search for longer cDNAs. In the Marathon™ technology, cDNAs have been prepared from mRNA extracted from a chosen tissue and an 'adaptor' sequence ligated onto each end. Nucleic acid amplification (PCR) is then carried out to amplify the "missing" 5' end of the DNA using a combination of gene specific and adaptor specific oligonucleotide primers. The PCR reaction is then repeated using "nested" primers, that is, primers designed to anneal within the amplified product (typically an adaptor specific primer that anneals further 3' in the adaptor sequence and a gene specific primer that anneals further 5' in the selected gene sequence). The products of this reaction can then be analyzed by DNA sequencing and a full-length DNA constructed either by joining the product directly to the existing DNA to give a complete sequence, or carrying out a separate full-length PCR using the new sequence information for the design of the 5' primer.

The polynucleotides and polypeptides of the invention may be employed, for example, as research reagents and materials for discovery of treatments of and diagnostics for diseases, particularly human diseases, as further discussed herein relating to polynucleotide assays.

The polynucleotides of the invention that are oligonucleotides derived from a sequence of SEQ Group 1 may be used in the processes herein as described, but preferably for PCR, to

determine whether or not the polynucleotides identified herein in whole or in part are transcribed in bacteria in infected tissue. It is recognized that such sequences will also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained.

5

The invention also provides polynucleotides that encode a polypeptide that is the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, may allow protein transport, may lengthen or shorten protein half-life or may facilitate manipulation of a protein for assay or production, among other things. As generally is the case *in vivo*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

15 For each and every polynucleotide of the invention there is provided a polynucleotide complementary to it. It is preferred that these complementary polynucleotides are fully complementary to each polynucleotide with which they are complementary.

A precursor protein, having a mature form of the polypeptide fused to one or more
20 prosequences may be an inactive form of the polypeptide. When prosequences are removed such inactive precursors generally are activated. Some or all of the prosequences may be removed before activation. Generally, such precursors are called proproteins.

In addition to the standard A, G, C, T/U representations for nucleotides, the term "N" may
25 also be used in describing certain polynucleotides of the invention. "N" means that any of the four DNA or RNA nucleotides may appear at such a designated position in the DNA or RNA sequence, except it is preferred that N is not a nucleic acid that when taken in combination with adjacent nucleotide positions, when read in the correct reading frame, would have the effect of generating a premature termination codon in such reading frame.

30

In sum, a polynucleotide of the invention may encode a mature protein, a mature protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences that are not the leader sequences of a preprotein, or a preproprotein, which is a precursor to a proprotein, having a leader sequence and one or more prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

In accordance with an aspect of the invention, there is provided the use of a polynucleotide of the invention for therapeutic or prophylactic purposes, in particular genetic immunization.

The use of a polynucleotide of the invention in genetic immunization will preferably employ a suitable delivery method such as direct injection of plasmid DNA into muscles (Wolff *et al.*, *Hum Mol Genet* (1992) 1: 363, Manthorpe *et al.*, *Hum. Gene Ther.* (1983) 4: 419), delivery of DNA complexed with specific protein carriers (Wu *et al.*, *J Biol Chem.* (1989) 264: 16985), coprecipitation of DNA with calcium phosphate (Benvenisty & Reshef, *PNAS USA*, (1986) 83: 9551), encapsulation of DNA in various forms of liposomes (Kaneda *et al.*, *Science* (1989) 243: 375), particle bombardment (Tang *et al.*, *Nature* (1992) 356:152, Eisenbraun *et al.*, *DNA Cell Biol* (1993) 12: 791) and *in vivo* infection using cloned retroviral vectors (Seeger *et al.*, *PNAS USA* (1984) 81: 5849).

Vectors, Host Cells, Expression Systems

The invention also relates to vectors that comprise a polynucleotide or polynucleotides of the invention, host cells that are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the invention.

Recombinant polypeptides of the present invention may be prepared by processes well known in those skilled in the art from genetically engineered host cells comprising expression systems. Accordingly, in a further aspect, the present invention relates to expression systems that comprise a polynucleotide or polynucleotides of the present invention, to host cells which are genetically engineered with such expression systems, and to the production of polypeptides of the invention by recombinant techniques.

For recombinant production of the polypeptides of the invention, host cells can be genetically engineered to incorporate expression systems or portions thereof or polynucleotides of the invention. Introduction of a polynucleotide into the host cell can be effected by methods described in many standard laboratory manuals, such as Davis, *et al.*, *BASIC METHODS IN MOLECULAR BIOLOGY*, (1986) and Sambrook, *et al.*, *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989), such as, calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, conjugation, transduction, scrape loading, ballistic introduction and infection.

Representative examples of appropriate hosts include bacterial cells, such as cells of streptococci, staphylococci, enterococci, *E. coli*, streptomyces, cyanobacteria, *Bacillus subtilis*, *Neisseria meningitidis*, *Haemophilus influenzae* and *Moraxella catarrhalis*; fungal cells, such as cells of a yeast, *Kluveromyces*, *Saccharomyces*, *Pichia*, a basidiomycete, *Candida albicans* and *Aspergillus*; insect cells such as cells of *Drosophila* S2 and *Spodoptera* Sf9; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, 293, CV-1 and Bowes melanoma cells; and plant cells, such as cells of a gymnosperm or angiosperm.

A great variety of expression systems can be used to produce the polypeptides of the invention. Such vectors include, among others, chromosomal-, episomal- and virus-derived vectors, for example, vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal

elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses, picornaviruses, retroviruses, and alphaviruses and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The

5 expression system constructs may contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to maintain, propagate or express polynucleotides and/or to express a polypeptide in a host may be used for expression in this regard. The appropriate DNA sequence may be inserted into the expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in
10 Sambrook *et al.*, *MOLECULAR CLONING, A LABORATORY MANUAL*, (*supra*).

In recombinant expression systems in eukaryotes, for secretion of a translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed
15 polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

Polypeptides of the present invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol
20 precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, ion metal affinity chromatography (IMAC) is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the
25 polypeptide is denatured during intracellular synthesis, isolation and or purification.

The expression system may also be a recombinant live microorganism, such as a virus or bacterium. The gene of interest can be inserted into the genome of a live recombinant virus or bacterium. Inoculation and *in vivo* infection with this live vector will lead to *in*
30 *vivo* expression of the antigen and induction of immune responses. Viruses and bacteria

used for this purpose are for instance: poxviruses (e.g. vaccinia, fowlpox, canarypox), alphaviruses (Sindbis virus, Semliki Forest Virus, Venezuelan Equine Encephalitis Virus), adenoviruses, adeno-associated virus, picornaviruses (poliovirus, rhinovirus), herpesviruses (varicella zoster virus, etc), *Listeria*, *Salmonella*, *Shigella*, BCG, streptococci. These viruses and bacteria can be virulent, or attenuated in various ways in order to obtain live vaccines. Such live vaccines also form part of the invention.

Combinations of Bordetella antigens in immunogenic compositions

- 10 A further aspect of the invention discloses particular combinations of Bordetella antigens which when combined, lead to an effective immunogenic composition against Bordetella infection. The efficacy of the immunogenic composition is determined as by its ability to elicit a protective response against *B. pertussis* primarily, but it is preferred that they also elicit a protective effect against the related bacteria *B. parapertussis* and/or *B. bronchiseptica*.

Preferred combinations of Bordetella antigens, when combined in an immunogenic composition or vaccine, allow different Bordetella functions to be targetted by the immune response. Such an immune response is better able to treat or prevent Bordetella infection. For instance, known virulence factors include adhesins like FHA, fimbriae, pertactin which are involved in attachment of Bordetella to host cells; toxins such as pertussis toxin, adenylate cyclase have a role in disabling the host immune system; BrkA acts as a serum resistance factor and TcfA has a role in tracheal colonization.

- 25 In particular, combinations of certain antigens from different classes, some of which are involved in adhesion to host cells, some of which are involved in iron acquisition, some of which are antotransporters and some of which are toxins, can elicit an immune response which protects against multiple functions of Bordetella required to sustain infection. Such combinations of antigens can surprisingly lead to synergistically

improved vaccine efficacy against *Bordetella* infection where more than one function of the bacterium is targeted by the immune response.

5 An additional advantage of preferred embodiments of the invention is that the combination of the antigens of the invention from different families of proteins in an immunogenic composition will enable protection against a wider range of strains. It is particularly advantageous for an immunogenic composition to generate an immune response against *B. pertussis* and *B. parapertussis* since *B. parapertussis* infections make up 5-10% of whooping cough cases and the protection afforded by currently used
10 acellular vaccines does not extend to *B. parapertussis*.

Accordingly, the invention provides immunogenic compositions comprising at least or exactly two, three, four, five, six, seven, eight, nine or ten different *Bordetella*, preferably *B. pertussis* antigens, wherein the antigens are selected from at least two,
15 three, four or five of the following categories:

- a) at least one *Bordetella* iron acquisition protein selected from the group consisting of the polypeptide sharing at least 70%, 80%, 90%, 95%, 97%, 98%, 99% or 100% identity with SEQ ID 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, or 28, or an
20 antigenic fragment thereof;
- b) at least one *Bordetella* autotransporter protein selected from the group consisting of pertactin, BipA, the polypeptide sharing at least 70%, 80%, 90%, 95%, 97%, 98%, 99% or 100% identity with SEQ ID 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, or 54, or an antigenic fragment thereof, preferably the passenger domain thereof;
- 25 c) at least one *Bordetella* lipoprotein selected from the group consisting of the polypeptide sharing at least 70%, 80%, 90%, 95%, 97%, 98%, 99% or 100% identity with SEQ ID 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, or 98 or an antigenic fragment thereof;
- d) at least one *Bordetella* adhesin selected from the group consisting of FHA, fimbriae,
30 pertactin and BrkA or an antigenic fragment thereof; and

e) at least one Bordetella toxin/invasin or antigens involved in toxin/invasin secretion selected from the group consisting of pertussis toxin, adenylate cyclase, dermonecrotic toxin (Dnt), Type III ss or lipopolysaccharide or an antigenic fragment thereof,

5 wherein the Bordetella antigens in the immunogenic composition do not consist of any combination of 2, 3, 4 or all 5 of pertactin, fimbriae 2, fimbriae 3, FHA and pertussis toxin.

10 Immunogenic compositions of the invention therefore do not cover known vaccines combinations, however a single known antigen from one group, combined with a new antigen from a different group is covered.

The Bordetella antigens may derived from any strain of Bordetella including from one or more of *B. pertussis*, *B. parapertussis* and *B. bronchiseptica* (preferably the former).

15 Preferably all five groups of antigen are represented in the immunogenic composition of the invention. Where an antigen falls into two groups, the inclusion of that one antigen into an immunogenic composition leads to the inclusion of both groups in the immunogenic composition.

20 Where a protein is specifically mentioned herein, it is preferably a reference to a native, full-length protein but it may also encompass antigenic, preferably immunogenic fragments thereof (particularly in the context of subunit vaccines). These are fragments containing or comprising at least 10 amino acids, preferably 20 amino acids, more
25 preferably 30 amino acids, more preferably 40 amino acids or most preferably 50 amino acids, taken contiguously from the amino acid sequence of the protein. In addition, antigenic fragments denotes fragments that are immunologically reactive with antibodies generated against the *B. pertussis* proteins or with antibodies generated by infection of a mammalian host with Bordetella. Antigenic fragments also includes
30 fragments that when administered at an effective dose, elicit a protective immune

response against *Bordetella* infection, more preferably it is protective against *B. pertussis* and/or *B. parapertussis* and/or *B. bronchiseptica* infection. Preferably such fragemtns are coupled to a source of T – cell epitopes.

- 5 Also included in the invention are recombinant fusion proteins of *Bordetella* proteins of the invention, or fragments thereof. These may combine different *Bordetella* proteins or fragments thereof in the same polypeptide. Alternatively, the invention also includes individual fusion proteins of *Bordetella* proteins or fragments thereof, as a fusion protein with heterologous sequences such as a provider of T-cell epitopes or purification
- 10 tags, for example: β -galactosidase, glutathione-S-transferase, green fluorescent proteins (GFP), epitope tags such as FLAG, myc tag, poly histidine, or viral surface proteins such as influenza virus haemagglutinin, tetanus toxoid, diphtheria toxoid, CRM197.

Antigens of the invention

1. *Autotransporter proteins*
- 15 Autotransporter proteins typically are made up of a signal sequence, a passenger domain and an anchoring domain for attachment to the outer membrane. Examples of autotransporter proteins include pertactin (SEQ ID 30), Vag8 (SEQ ID 32), BrkA (SEQ ID 34), TcfA (SEQ ID 36) (Finn and Stevens (1995) Mol. Microbiol. 16; 625-634), Phg (SEQ ID 38), BipA (Stockbauer et al 2001; Molecular Microbiology 39; 65-78), BapA
- 20 (SEQ ID 40), BapB (SEQ ID 42), BapC (SEQ ID 44), pertactin-like protein (SEQ ID 46), Tcf-like protein (SEQ ID 48), extracellular serine protease (SEQ ID 50), YapE (SEQ ID 52), SphB1 (SEQ ID 54). These antigens may be derived from *Bordetella pertussis* or *Bordetella parapertussis* or other *Bordetella* strains.
- 25 It is particularly advantageous to use the passenger domain of an autotranporter when it is included in a subunit vaccine. Table 2 above defines the passsenger domains of the autotransporter proteins listed above.

BipA contains 90 amino acid tandem repeats with 5 being present in the *B. pertussis* protein and 8 being present in the *B. bronchiseptica* protein. These repeats span from amino acid 581 to 1030 in *B. pertussis* and amino acids 581 to 1300 in *B.*

bronchiseptica. Preferred fragments of BipA include amino acids 1031 to 1308, amino acids 941 to 1308, amino acids 851 to 1308, amino acids 761 to 1308, amino acids 671 to 1308 and 581 to 1308 of the *B. pertussis* sequence (or sequences related to these that have 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40 or 50 amino acids added or deleted from either or both of the N and C termini).

10 2. *Iron acquisition proteins*

Iron acquisition is of great importance to mammalian pathogens as iron is present in limiting conditions in the host and any iron that is present is sequestered by haem and other iron-chelating compounds. Iron from siderophores and host iron-binding complexes is internalised through TonB-dependent outer membrane ferric complex receptors. *Bordetella* iron acquisition proteins include BfeA (SEQ ID 2), BfrB (SEQ ID 4), BfrC (SEQ ID 6), FauA (SEQ ID 8), ferric siderophore receptor (SEQ ID 10), Ferric alcaligin siderophore receptor (SEQ ID 12), iron transport protein fiu (SEQ ID 14), iron transport protein fiu (SEQ ID 16), putative hydroxamate-type ferrisiderophore receptor signal peptide protein (SEQ ID 18) BhuR (SEQ ID 20) (Infection and Immunity 2001, 69; 6951), tonb-dependent receptor (SEQ ID 22), tonb-dependent receptor (SEQ ID 24), ferrisiderophore receptor-like protein (SEQ ID 26) and tonb-dependent receptor Yncd precursor (SEQ ID 28). These proteins may be derived from *Bordetella pertussis*, *Bordetella parapertussis* or other *Bordetella* strains, preferably the former.

25 3. *Lipoproteins*

~~*Bordetella* lipoproteins include heme/hemopexin-utilisation protein C precursor (SEQ ID 56), piln protein (SEQ ID 58), immunogenic protein (SEQ ID 60), outer membrane lipoprotein precursor (SEQ ID 62), outer membrane efflux protein precursor (SEQ ID~~

64), Oprm (SEQ ID 66), outer membrane channel signal protein (SEQ ID 68), MltA (SEQ ID 70), MltB (SEQ ID 72), yccz precursor (SEQ ID 74), serine protease transmembrane protein (SEQ ID 76), pa4632 (SEQ ID 78), coml precursor (SEQ ID 80), VacJ (SEQ ID 82), outer membrane lipoprotein (SEQ ID 84), Flagellar 1-ring protein (SEQ ID 86), YdcI (SEQ ID 88), Pal (SEQ ID 90), OmlA (SEQ ID 92), Smc00354 (SEQ ID 94), Pcp (SEQ ID 96) and lipoprotein (SEQ ID 98).

The lipoproteins having the sequence of SEQ ID 56-96 contain a lipidation motif indicating that they would be lipidated and inserted into the membrane. In its simplest form, the lipidation motif contains the consensus sequence LXXC. However, the consensus sequence is preferably close to the amino terminus of the sequence, within the larger consensus sequence:-

<(M, V, L)~*{1, 40}~(D, E, R, K,*)6(L, I, V, M, F, W, S, T, A, G)2(L, I, V, M, F, Y, S, T, A, G, C, Q) (A, G, S) C

< indicates the amino terminus of the protein so that the first amino acid should be M, V or L. {1, 40} indicates that between 1 and 40 amino acids should be present between the first amino acid and the rest of the consensus sequence. (D, E, R, K, *)6 indicates that the next 6 amino acids should not be D, E, R or K. The following 2 amino acids should be one of the aliphatic amino acids indicated and is preferably L. For the following 2 amino acids, the amino acids shown in the parentheses should be present and the final amino acid of the sequence should be C. These antigens may be derived from *Bordetella pertussis* or *Bordetella parapertussis*, *Bordetella bronchiseptica* or other *Bordetella* strains, preferably the former.

4. Adhesins

Adhesins have a role in attaching *Bordetella* to a host cell and hence have important roles in virulence. They include filamentous haemagglutinin (FHA) (Relman et al (1989) Proc Natl Acad Sci USA 86; 2634-2641), fimbriae (Fim) (Mooi et al (1992)

Microb Pathog 12; 127-135), pertactin (Roberts et al (1991) Mol Microbiol 5; 1393-1404) and BrkA (Fernandez et al (1994) Infection and Immunity 62; 4727-4738).

These antigens may be derived from *Bordetella pertussis* or *Bordetella parapertussis*, *Bordetella bronchoseptica* or other *Bordetella* strains, preferably the former.

5

5. Toxins

Toxins include adenylate cyclase (CyaA) (Hewlett et al (1989) J. Biol.Chem. 264; 19379-19384), pertussis toxin (Munoz et al (1981) Infect Immun 33; 820-826), dermo-
 10 necrotic toxin (Dnt) (Livey (1984) J. Med. Microbiol. 17; 91-103 and lipopolysaccharides. Toxins also include proteins that are involved in the secretion of toxins since an immune response against the secretory mechanism would prevent the efficient functioning of the secretory mechanism and lead to reduced toxin secretion. An example of such an antigen is the Type III secretion system (Yuk et al (2000) Mol.
 15 Microbiol. 35; 991-1004). These antigens may be derived from *Bordetella pertussis* or *Bordetella parapertussis*, *Bordetella bronchoseptica* or other *Bordetella* strains, preferably the former.

Where toxin is described herein, non-toxic derivatives such as toxoids or mutant toxins
 20 are also envisaged to be covered by the term.

Preferred combinations

In a preferred embodiment of the combination of antigens of the invention, the
 25 immunogenic composition comprises at least one *Bordetella* iron acquisition protein selected from the group consisting of the polypeptide sharing at least 70%, 80%, 90%, 95%, 97%, 98%, 99% or 100% amino acid identity with SEQ ID 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, or an antigenic fragment thereof.

In a particularly preferred embodiment of the invention, the immunogenic composition comprises one, two or three of FHA, pertussis toxin and pertactin, (preferably FHA and PT; FHA and pertactin; PT and pertactin; or FHA, pertussis toxin and pertactin) and further comprises at least one Bordetella iron acquisition protein selected from the group consisting of the polypeptide sharing at least 70%, 80%, 90%, 95%, 97%, 98%, 99% or 100% amino acid identity with SEQ ID 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, or an antigenic fragment thereof.

In a further preferred embodiment of the invention, the immunogenic composition comprises at least one Bordetella autotransporter protein selected from the group consisting of BipA, the polypeptide sharing at least 70%, 80%, 90%, 95%, 97%, 98%, 99% or 100% amino acid identity with SEQ ID 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, or an antigenic fragment thereof, preferably the passenger domain thereof.

In a particularly preferred embodiment of the invention, the immunogenic composition comprises one, two or three of FHA, pertussis toxin and pertactin, (preferably FHA and PT; FHA and pertactin; PT and pertactin; or FHA, pertussis toxin and pertactin), and further comprises at least one Bordetella autotransporter protein selected from the group consisting of BipA, the polypeptide sharing at least 70%, 80%, 90%, 95%, 97%, 98%, 99% or 100% amino acid identity with SEQ ID 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, or an antigenic fragment thereof, preferably the passenger domain thereof.

In a further preferred embodiment of the invention, the immunogenic composition will comprise at least one Bordetella lipoprotein selected from the group consisting of the polypeptide sharing at least 70%, 80%, 90%, 95%, 97%, 98%, 99% amino acid identity with SEQ ID 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98 or an antigenic fragment thereof.

In a particularly preferred embodiment of the invention, the immunogenic composition comprises one, two or three of FHA, pertussis toxin and pertactin, (preferably FHA and

PT; FHA and pertactin; PT and pertactin; or FHA, pertussis toxin and pertactin) at least one Bordetella lipoprotein selected from the group consisting of the polypeptide sharing at least 70%, 80%, 90%, 95%, 97%, 98%, 99% amino acid identity with SEQ ID 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98 or an
5 antigenic fragment thereof.

In a further preferred embodiment of the invention, the immunogenic composition comprises BrkA or an antigenic fragment thereof.

- 10 In a further preferred embodiment of the invention, the immunogenic composition comprises at least one Bordetella toxin or antigens involved in toxin secretion selected from the group consisting of adenylate cyclase, dermonecrotic toxin (Dnt), Type III ss or lipopolysaccharide or an antigenic fragment thereof.
- 15 The combination of FHA, pertussis toxin and pertactin is known to elicit a protective immune response and particularly preferred combinations of the invention will contain one, two or three of these constituents, optionally further comprising fim 2, fim 3 or fim 2 and fim 3.
- 20 A preferred immunogenic composition of the invention contains TcfA and at least 1, 2, 3, 4, antigens selected from the list consisting of pertussis toxin, adenylate cyclase, LPS, BipA, Type III ss, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp. Preferred combinations comprise TcfA and pertussis toxin, (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp);
25 TcfA and adenylate cyclase (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); TcfA and LPS (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); TcfA and BipA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate

cyclase, LPS, Type IIIss, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA
30 and Pcp); TcfA and Type IIIss (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, BipA,

- BhuR , FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); TcfA and BhuR (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA, PCP); TcfA and FHA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, pertactin, BhuR, Type IIIss, MltA, MltB, VacJ, OmlA, Pcp); TcfA and a lipoprotein selected from MltA, MltB, VacJ, OmlA and Pcp (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, BhuR, Type IIIss, Fim, pertactin, BrkA and FHA).
- 10 A further preferred immunogenic composition of the invention contains BipA and at least 1, 2, 3, 4, antigens selected from the list consisting of pertussis toxin, adenylate cyclase, LPS, Type III ss, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp. Preferred combinations comprise BipA and pertussis toxin, (optionally with 1, 2, 3, 4 or 5 of BhuR, FHA, pertactin, Fim, BrkA, MltA, MltB, VacJ, OmlA and Pcp),
- 15 BipA and adenylate cyclase (optionally with 1, 2, 3, 4 or 5 of BhuR, FHA, pertactin, Fim, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BipA and LPS (optionally with 1, 2, 3, 4 or 5 of BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BipA and Type IIIss (optionally with 1, 2, 3, 4 or 5 of BhuR , FHA, Fim, pertactin, BrkA, pertussis toxin, MltA, MltB, VacJ, OmlA and Pcp); BipA and BhuR (optionally with 1,
- 20 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, Type IIIss, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and PCP); BipA and FHA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, pertactin, BhuR, Type IIIss, MltA, MltB, VacJ, OmlA and PCP); BipA and a lipoprotein selected from MltA, MltB, VacJ, OmlA and Pcp (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase,
- 25 LPS, BhuR, Type IIIss, Fim, pertactin, BrkA and FHA).

- A further preferred immunogenic composition of the invention contains BapA and at least 1, 2, 3, 4, antigens selected from the list consisting of pertussis toxin, adenylate cyclase, LPS, BipA, Type III ss, BhuR, FHA, Fim, pertactin, BrkA , MltA, MltB, VacJ,
- 30 OmlA and Pcp. Preferred combinations comprise BapA and pertussis toxin, (optionally

with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BapA and adenylate cyclase (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, pertussis toxin, MltA, MltB, VacJ, OmlA and Pcp); BapA and LPS (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BapA and BipA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, Type IIIss, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BapA and Type IIIss (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BapA and BhuR (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BapA and FHA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, BhuR, Fim, pertactin, BrkA, Type IIIss, MltA, MltB, VacJ, OmlA and Pcp); BapA and a lipoprotein selected from MltA, MltB, VacJ, OmlA and Pcp (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Fim, pertactin, BrkA, BhuR, Type IIIss and FHA).

A further preferred immunogenic composition of the invention contains BapB and at least 1, 2, 3, 4, antigens selected from the list consisting of pertussis toxin, adenylate cyclase, LPS, BipA, Type III ss, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp. Preferred combinations comprise BapB and pertussis toxin, (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BapB and adenylate cyclase (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BapB and LPS (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BapB and BipA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, Type IIIss, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BapB and Type IIIss (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BapB and BhuR (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA

and Pcp); BapB and FHA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Fim, pertactin, BrkA, BhuR, Type IIIss, MltA, MltB, VacJ, OmlA and Pcp); BapB and a lipoprotein selected from MltA, MltB, VacJ, OmlA and Pcp (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, BhuR, Type IIIss, Fim, pertactin, BrkA and FHA).

A further preferred immunogenic composition of the invention contains BapC and at least 1, 2, 3, 4, antigens selected from the list consisting of pertussis toxin, adenylate cyclase, LPS, BipA, Type III ss, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and PCP. Preferred combinations comprise BapC and pertussis toxin, (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BapC and adenylate cyclase (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BapC and LPS (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BapC and BipA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, Type IIIss, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BapC and Type IIIss (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BapC and BhuR (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BapC and FHA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, BhuR, Type IIIss, pertactin, MltA, MltB, VacJ, OmlA and Pcp); BapC and a lipoprotein selected from MltA, MltB, VacJ, OmlA and PCP (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, BhuR, Type IIIss, Fim, pertactin, BrkA and FHA).

A further preferred immunogenic composition of the invention contains pertactin and at least 1, 2, 3, 4, antigens selected from the list consisting of pertussis toxin, adenylate cyclase, LPS, BipA, Type III ss, BhuR, FHA, Fim, BrkA, MltA, MltB, VacJ, OmlA and PCP. Preferred combinations comprise pertactin and pertussis toxin, (optionally with 1,

2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, BrkA, MltA, MltB, VacJ, OmlA and Pcp);
 pertactin and adenylate cyclase (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA,
 Fim, BrkA, MltA, MltB, VacJ, OmlA and Pcp), pertactin and LPS (optionally with 1,
 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, BrkA, MltA, MltB, VacJ, OmlA and Pcp);
 5 pertactin and BipA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase,
 LPS, Type IIIss, BhuR, FHA, Fim, BrkA, MltA, MltB, VacJ, OmlA and Pcp); pertactin
 and Type IIIss (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, BipA, BhuR, FHA,
 Fim, BrkA, MltA, MltB, VacJ, OmlA and Pcp); pertactin and BhuR (optionally with 1,
 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim,
 10 BrkA, MltA, MltB, VacJ, OmlA and Pcp); pertactin and FHA (optionally with 1, 2, 3,
 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, BhuR, Type IIIss, MltA, MltB,
 VacJ, OmlA and Pcp), pertactin and a lipoprotein selected from MltA, MltB, VacJ,
 OmlA and PCP (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase,
 LPS, BipA, BhuR, Type IIIss, Fim, BrkA and FHA).

15

A further preferred immunogenic composition of the invention contains pertactin-like
 protein and at least 1, 2, 3, 4, antigens selected from the list consisting of pertussis
 toxin, adenylate cyclase, LPS, BipA, Type III ss, BhuR, FHA, pertactin, Fim, BrkA,
 MltA, MltB, VacJ, OmlA and Pcp. Preferred combinations comprise pertactin-like
 20 protein and pertussis toxin, (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA,
 pertactin, Fim, BrkA, MltA, MltB, VacJ, OmlA and Pcp); pertactin-like protein and
 adenylate cyclase (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, pertactin, Fim,
 BrkA, MltA, MltB, VacJ, OmlA and Pcp); pertactin-like protein and LPS (optionally
 with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, pertactin, Fim, BrkA, MltA, MltB, VacJ,
 25 OmlA and PCP); pertactin-like protein and BipA (optionally with 1, 2, 3, 4 or 5 of
 pertussis toxin, adenylate cyclase, LPS, Type IIIss, BhuR, FHA, pertactin, Fim, BrkA,
 MltA, MltB, VacJ, OmlA and Pcp); pertactin-like protein and Type IIIss (optionally
 with 1, 2, 3, 4 or 5 of pertussis toxin, BipA, BhuR, FHA, pertactin, Fim, BrkA, MltA,
 MltB, VacJ, OmlA and Pcp); pertactin-like protein and BhuR (optionally with 1, 2, 3, 4
 30 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, pertactin, Fim,

BrkA, MltA, MltB, VacJ, OmlA, Pcp); pertactin-like protein and FHA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, BhuR, pertactin, Fim, BrkA, Type IIIss, MltA, MltB, VacJ, OmlA, Pcp); pertactin-like protein and a lipoprotein selected from MltA, MltB, VacJ, OmlA and Pcp (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, BhuR, Type IIIss, pertactin, Fim, BrkA and FHA).

A further preferred immunogenic composition of the invention contains YapE and at least 1, 2, 3, 4, antigens selected from the list consisting of pertussis toxin, adenylate cyclase, LPS, BipA, Type III ss, BhuR, FHA, pertactin, Fim, BrkA, MltA, MltB, VacJ, OmlA, Pcp. Preferred combinations comprise YapE and pertussis toxin, (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, pertactin, Fim, BrkA, MltA, MltB, VacJ, OmlA, Pcp); YapE and adenylate cyclase (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, pertactin, Fim, BrkA, MltA, MltB, VacJ, OmlA, Pcp); YapE and LPS (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, pertactin, Fim, BrkA, MltA, MltB, VacJ, OmlA, Pcp); YapE and BipA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, Type IIIss, BhuR, FHA, pertactin, Fim, BrkA, MltA, MltB, VacJ, OmlA, Pcp); YapE and Type IIIss (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, BipA, BhuR, FHA, pertactin, Fim, BrkA, MltA, MltB, VacJ, OmlA, Pcp); YapE and BhuR (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, pertactin, Fim, BrkA, MltA, MltB, VacJ, OmlA, Pcp); YapE and FHA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, BhuR, Type IIIss, pertactin, MltA, MltB, VacJ, OmlA, Pcp); YapE and a lipoprotein selected from MltA, MltB, VacJ, OmlA, Pcp (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, BhuR, Type IIIss, pertactin, Fim, BrkA and FHA).

A further preferred immunogenic composition of the invention contains BrkA and at least 1, 2, 3, 4, antigens selected from the list consisting of pertussis toxin, adenylate cyclase, LPS, BipA, Type III ss, BhuR, FHA, pertactin, Fim, MltA, MltB, VacJ, OmlA,

Pcp. Preferred combinations comprise BrkA and pertussis toxin, (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, pertactin, Fim, MltA, MltB, VacJ, OmlA, Pcp); BrkA and adenylate cyclase (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, pertactin, Fim, MltA, MltB, VacJ, OmlA, Pcp); BrkA and LPS (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, pertactin, Fim, MltA, MltB, VacJ, OmlA and Pcp); BrkA and BipA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, Type IIIss, BhuR, FHA, pertactin, Fim, MltA, MltB, VacJ, OmlA, Pcp), BrkA and Type IIIss (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, BipA, BhuR, FHA, pertactin, Fim, MltA, MltB, VacJ, OmlA, Pcp); BrkA and BhuR (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, pertactin, Fim, MltA, MltB, VacJ, OmlA, Pcp); BrkA and FHA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, BhuR, Type IIIss, pertactin, MltA, MltB, VacJ, OmlA, Pcp); BrkA and a lipoprotein selected from MltA, MltB, VacJ, OmlA and Pcp (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, BhuR, Type IIIss, pertactin, Fim and FHA).

A further preferred immunogenic composition of the invention contains FHA and at least 1, 2, 3, 4, antigens selected from the list consisting of pertussis toxin, adenylate cyclase, LPS, BipA, Type III ss, BhuR, BrkA pertactin, Fim, MltA, MltB, VacJ, OmlA and Pcp. Preferred combinations comprise FHA and pertussis toxin, (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, BrkA, pertactin, Fim, MltA, MltB, VacJ, OmlA and Pcp); FHA and adenylate cyclase (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, BrkA, pertactin, Fim, MltA, MltB, VacJ, OmlA and Pcp); FHA and LPS (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, BrkA, pertactin, Fim, MltA, MltB, VacJ, OmlA and Pcp); FHA and BipA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, Type IIIss, BhuR, BrkA, pertactin, Fim, MltA, MltB, VacJ, OmlA and Pcp); FHA and Type IIIss (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, BipA, BhuR, BrkA, pertactin, Fim, MltA, MltB, VacJ, OmlA and Pcp); FHA and BhuR (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, BrkA, pertactin, Fim, MltA, MltB, VacJ, OmlA and Pcp); FHA and a lipoprotein selected

from MltA, MltB, VacJ, OmlA and Pcp (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, BhuR, Type IIIss, pertactin, Fim and BrkA).

A further preferred immunogenic composition of the invention contains BhuR and at least 1, 2, 3, 4, antigens selected from the list consisting of pertussis toxin, adenylate cyclase, LPS, TcfA, BapA, BapB, BapC, pertactin, pertactin-like protein, YapE, BrkA, BipA, Type III ss, FHA, Fim, MltA, MltB, VacJ, OmlA and Pcp. Preferred combinations comprise BhuR and pertussis toxin (optionally with 1, 2, 3, 4 or 5 of TcfA, BapA, BapB, BapC, pertactin, pertactin-like protein, YapE, BrkA, BipA, FHA, Fim, MltA, MltB, VacJ, OmlA and Pcp); BhuR and adenylate cyclase (optionally with 1, 2, 3, 4 or 5 of TcfA, BapA, BapB, BapC, pertactin, pertactin-like protein, YapE, BrkA, BipA, pertussis toxin, FHA, Fim, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BhuR and LPS (optionally with 1, 2, 3, 4 or 5 of TcfA, BapA, BapB, BapC, pertactin, pertactin-like protein, YapE, BrkA, BipA, pertussis toxin, FHA, Fim, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BhuR and TcfA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BhuR and BapA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BhuR and BapB (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BhuR and BapC (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BhuR and pertactin (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BhuR and pertactin-like protein (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BhuR and YapE (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BhuR and BrkA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim,

pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); BhuR and a lipoprotein selected from MltA, MltB, VacJ, OmlA and Pcp (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, TcfA, BapA, BapB, BapC, pertactin, pertactin-like protein, YapE, BrkA, BipA, Type IIIss, Fim and FHA).

5

A further preferred immunogenic composition of the invention contains MltA and at least 1, 2, 3, 4, antigens selected from the list consisting of pertussis toxin, adenylate cyclase, LPS, TcfA, BapA, BapB, BapC, pertactin, pertactin-like protein, YapE, BrkA, BipA, Type III ss, BhuR, Fim and FHA. Preferred combinations comprise MltA, and
 10 pertussis toxin (optionally with 1, 2, 3, 4 or 5 of TcfA, BapA, BapB, BapC, pertactin, pertactin-like protein, YapE, BrkA, BipA, FHA, Fim and BhuR); MltA and adenylate cyclase (optionally with 1, 2, 3, 4 or 5 of TcfA, BapA, BapB, BapC, pertactin, pertactin-like protein, YapE, BrkA, BipA, pertussis toxin, FHA, Fim and BhuR); MltA and LPS (optionally with 1, 2, 3, 4 or 5 of TcfA, BapA, BapB, BapC, pertactin,
 15 pertactin-like protein, YapE, BrkA, BipA, pertussis toxin, FHA, Fim and BhuR); MltA and TcfA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, pertactin, BrkA and BhuR); MltA and BapA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, pertactin, BrkA and BhuR); MltA and BapB (optionally with 1, 2, 3, 4 or 5 of
 20 pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, pertactin, BrkA and BhuR); MltA and BapC (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, pertactin, BrkA and BhuR); MltA and pertactin (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, BrkA and BhuR); MltA and pertactin-like protein
 25 (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, pertactin, BrkA and BhuR); MltA and YapE (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, pertactin, BrkA and BhuR); MltA and BrkA (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, BipA, Type IIIss, FHA, Fim, pertactin and BhuR); MltA and
 30 BhuR (optionally with 1, 2, 3, 4 or 5 of pertussis toxin, adenylate cyclase, LPS, TcfA,

BapA, BapB, BapC, pertacin, pertactin-like protein, YapE, BrkA, BipA, Type IIIss, Fim and FHA).

A further preferred immunogenic composition of the invention contains pertussis toxin and at least 1, 2, 3, 4, antigens selected from the list consisting of TcfA, BapA, BapB, BapC, pertactin, pertactin-like protein, YapE, BrkA, BipA, BhuR, FHA, MltA, MltB, VacJ, OmlA and Pcp. Preferred combinations comprise pertussis toxin and TcfA (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); pertussis toxin and BapA (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); pertussis toxin and BapB (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); pertussis toxin and BapC (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); pertussis toxin and pertactin (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, BrkA, MltA, MltB, VacJ, OmlA and Pcp); pertussis toxin and pertactin-like protein (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); pertussis toxin and YapE (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, MltA, MltB, VacJ, OmlA and Pcp); pertussis toxin and BrkA (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, MltA, MltB, VacJ, OmlA and Pcp); pertussis toxin and BhuR (optionally with 1, 2, 3, 4 or 5 of TcfA, BapA, BapB, BapC, pertactin, pertactin-like protein, YapE, BrkA, BipA, FHA, Fim, MltA, MltB, VacJ, OmlA and Pcp); pertussis toxin and MltA, MltB (optionally with 1, 2, 3, 4 or 5 of TcfA, BapA, BapB, BapC, pertactin, pertactin-like protein, YapE, BrkA, BipA, FHA, Fim and BhuR).

25

A further preferred immunogenic composition of the invention contains adenylate cyclase and at least 1, 2, 3, 4, antigens selected from the list consisting of TcfA, BapA, BapB, BapC, pertactin, pertactin-like protein, YapE, BrkA, BipA, BhuR, FHA, Fim, pertussis toxin, MltA, MltB, VacJ, OmlA and Pcp. Preferred combinations comprise adenylate cyclase and TcfA (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim,

30

pertactin, BrkA, pertussis toxin , MltA, MltB, VacJ, OmlA and Pcp); adenylate cyclase and BapA (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, pertussis toxin , MltA, MltB, VacJ, OmlA and Pcp); adenylate cyclase and BapB (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, pertussis toxin , MltA, MltB, VacJ, OmlA and Pcp); adenylate cyclase and BapC (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, pertussis toxin , MltA, MltB, VacJ, OmlA and Pcp); adenylate cyclase and pertactin (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, pertussis toxin , MltA, MltB, VacJ, OmlA and Pcp); adenylate cyclase and pertactin-like protein (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, pertussis toxin , MltA, MltB, VacJ, OmlA and Pcp); adenylate cyclase and YapE (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, pertussis toxin , MltA, MltB, VacJ, OmlA and Pcp); adenylate cyclase and BrkA (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, pertussis toxin , MltA, MltB, VacJ, OmlA and Pcp), adenylate cyclase and BhuR (optionally with 1, 2, 3, 4 or 5 of TcfA, BapA, BapB, BapC, pertactin, pertactin-like protein, YapE, BrkA, BipA, FHA, Fim pertussis toxin , MltA, MltB, VacJ, OmlA and Pcp); adenylate cyclase and MltA, MltB (optionally with 1, 2, 3, 4 or 5 of TcfA, BapA, BapB, BapC, pertactin, pertactin-like protein, YapE, BrkA, BipA, FHA, Fim, pertussis toxin and BhuR).

20

A further preferred immunogenic composition of the invention contains LPS and at least 1, 2, 3, 4, antigens selected from the list consisting of TcfA, BapA, BapB, BapC, pertactin, pertactin-like protein, YapE, BrkA, BipA, BhuR, FHA, Fim, pertussis toxin , MltA, MltB, VacJ, OmlA and Pcp. Preferred combinations comprise LPS and TcfA (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, pertussis toxin , MltA, MltB, VacJ, OmlA and Pcp); LPS and BapA (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, pertussis toxin , MltA, MltB, VacJ, OmlA and Pcp); LPS and BapB (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, pertussis toxin , MltA, MltB, VacJ, OmlA and Pcp); LPS and BapC (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA,

30

pertussis toxin , MltA, MltB, VacJ, OmlA and Pcp); LPS and pertactin (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, BrkA, pertussis toxin , MltA, MltB, VacJ, OmlA and Pcp); LPS and pertactin-like protein (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, pertussis toxin , MltA, MltB, VacJ, OmlA and Pcp);

5 LPS and YapE (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, BrkA, pertussis toxin , MltA, MltB, VacJ, OmlA and Pcp); LPS and BrkA (optionally with 1, 2, 3, 4 or 5 of BipA, BhuR, FHA, Fim, pertactin, pertussis toxin , MltA, MltB, VacJ, OmlA and Pcp); LPS and BhuR (optionally with 1, 2, 3, 4 or 5 of TcfA, BapA, BapB, BapC, pertactin, pertactin-like protein, YapE, BrkA, BipA, FHA, Fim, pertussis

10 toxin , MltA, MltB, VacJ, OmlA and Pcp); LPS and MltA (optionally with 1, 2, 3, 4 or 5 of TcfA, BapA, BapB, BapC, pertactin, pertactin-like protein, YapE, BrkA, BipA, FHA, Fim and pertussis toxin and BhuR).

A further preferred combination of the invention contains one, two or three of FHA,

15 pertactin, pertussis toxin (preferably FHA and pertactin, FHA and pertussis toxin, pertactin and pertussis toxin or FHA, pertactin and pertussis toxin) and an additional 1, 2, 3 or 4 antigens selected from the group consisting of TcfA, BipA, BapA, BapB, BapC, pertactin-like protein, YapE, BhuR, Fim, BrkA, adenylate cyclase, Type III ss , MltA, MltB, VacJ, OmlA and Pcp. A preferred combinations contains FHA, pertactin, pertussis

20 toxin and BhuR. A further preferred combination contains FHA, pertactin, pertussis toxin , and MltA. A further preferred combination contains FHA, pertactin, pertussis toxin , and MltB. A further preferred combination contains FHA, pertactin, pertussis toxin and VacJ. A further preferred combination contains FHA, pertactin, pertussis toxin and OmlA. A further preferred combination contains FHA, pertactin, pertussis toxin and Pcp. A

25 further preferred combination contains FHA, pertactin, pertussis toxin, a lipoprotein (preferably MltA, MltB, VacJ, OmlA or Pcp) and BhuR.

A preferred immunogenic composition of the invention comprises FHA, pertussis toxin and BrkA or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with

30 SEQ ID 34, preferably further comprising pertactin.

A preferred immunogenic composition of the invention comprises FHA, pertussis toxin and BhuR or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 20, preferably further comprising pertactin.

5

A preferred immunogenic composition of the invention comprises FHA, pertussis toxin and BapA or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 40, preferably further comprising pertactin.

10 A preferred immunogenic composition of the invention comprises FHA, pertussis toxin and BapB or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 42, preferably further comprising pertactin.

15 A preferred immunogenic composition of the invention comprises FHA, pertussis toxin and BapC or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 44, preferably further comprising pertactin.

A preferred immunogenic composition of the invention comprises FHA, pertussis toxin and YapE, preferably further comprising pertactin.

20

A preferred immunogenic composition of the invention comprises FHA, pertussis toxin and VacJ or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 82, preferably further comprising pertactin.

25 A preferred immunogenic composition of the invention comprises FHA, pertussis toxin and OmlA or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 92, preferably further comprising pertactin.

A preferred immunogenic composition of the invention comprises FHA, pertussis toxin and Pcp or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 96, preferably further comprising pertactin.

- 5 A preferred immunogenic composition of the invention comprises FHA, pertussis toxin and MltA or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 70, preferably further comprising pertactin.

- 10 A preferred immunogenic composition of the invention comprises FHA, pertussis toxin and MltB or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 72, preferably further comprising pertactin.

- 15 A preferred immunogenic composition of the invention comprises FHA, pertussis toxin and TcfA or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 36, preferably further comprising pertactin.

A preferred immunogenic composition of the invention comprises FHA, pertussis toxin and adenylate cyclase, preferably further comprising pertactin.

- 20 A preferred immunogenic composition of the invention comprises FHA, pertussis toxin and Type III ss, preferably further comprising pertactin.

- 25 It is further advantageous to combine antigens that are present at different stages of Bordetella life cycle in the immunogenic compositions of the invention. Bordetella has three identifiable stages, during which protein expression is controlled by the bvgAS locus. The Bvg⁺ virulent phase is characterised by the expression of a number of virulence factors including FHA, fimbriae and pertactin, a variety of toxins including adenylate cyclase, dermonecrotic toxin and pertussis toxin. During the Bvgⁱ phase, some of the virulence factors are expressed and a new set of proteins including BipA are
30 expressed (Deora et al Molecular Microbiology (2001) 40; 669-683).

Currently available acellular pertussis vaccines include FHA, adenylate cyclase, pertactin and fimbriae proteins which are all Bvg⁺ early genes. Accordingly, a further aspect of the invention is an immunogenic composition containing 2 or more antigens which are expressed in two, three or four phases selected from Bvg⁺ early, Bvg⁺ late, Bvg⁻ and Bvgi, for example, Bvg⁺ early and Bvg⁺late; Bvg⁺ early and Bvg⁻; Bvg⁺ early and Bvgi; Bvg⁺late and Bvg⁻; Bvg⁺ late and Bvgi; Bvg⁻ and Bvgi; Bvg⁺ early and Bvg⁺late and Bvg⁻; Bvg⁺ early and Bvg⁺late and Bvgi; Bvg⁺ early and Bvg⁻ and Bvgi; Bvg⁺ late and Bvg⁻ and Bvgi; Bvg⁺ early and Bvg⁺late and Bvg⁻ and Bvgi.

10

FHA, pertussis toxin, adenylate cyclase, Fim and pertactin are expressed in Bvg⁺ early phase.

Vag8, SphB1, Tcf and Type III SS are expressed in Bvg⁺ late phase.

BipA is expressed during Bvgi phase.

15 LPS are present in all phases including Bvg⁻ phase.

Accordingly, preferred immunogenic compositions of the invention comprise 1, 2 or 3 antigens expressed in Bvg⁺ early phase (preferably selected from FHA, pertussis toxin, Fim and pertactin and further comprise 1, 2 or 3 antigens that are expressed during Bvg⁺ late phase and/or Bvgi phase and/or Bvg⁻ phase (preferably Bvgi).

20

A preferred immunogenic compositions comprises FHA, PT and Tcf (optionally further comprising 1, 2, or 3 of Fim, pertactin, Vag8, SphB1, Type III SS, BipA and LPS).

25 A preferred immunogenic compositions comprises FHA, PT and Vag8 (optionally further comprising 1, 2, or 3 of Fim, pertactin, Tcf, SphB1, Type III SS, BipA and LPS).

A preferred immunogenic compositions comprises FHA, PT and Vag8 (optionally further comprising 1, 2, or 3 of Fim, pertactin, Tcf, SphB1, Type III SS, BipA and LPS).

30

A preferred immunogenic compositions comprises FHA, PT and SphB1 (optionally further comprising 1, 2, or 3 of Fim, pertactin, Tcf, Vag8, Type III SS, BipA and LPS).

- 5 A preferred immunogenic compositions comprises FHA, PT and Type III SS (optionally further comprising 1, 2, or 3 of Fim, pertactin, Tcf, Vag8, SphB1, BipA and LPS).

A preferred immunogenic compositions comprises FHA, PT and BipA (optionally further comprising 1, 2, or 3 of Fim, pertactin, Tcf, Vag8, SphB1, Type III SS, BipA and
10 LPS).

A preferred immunogenic compositions comprises FHA, PT and LPS (optionally further comprising 1, 2, or 3 of Fim, pertactin, Tcf, Vag8, SphB1, Type III SS and BipA).

- 15 The combinations listed above may be in the form of a subunit vaccine which contains isolated, preferably purified antigens. Where this is the case, it is preferred that soluble fragments of some of the antigens are used. For instance, the water soluble passenger domain of autotransporter proteins as defined above are preferred.

20 **Diagnostic, Prognostic, Serotyping and Mutation Assays**

This invention is also related to the use of BASB232 polynucleotides and polypeptides of the invention for use as diagnostic reagents. Detection of BASB232 polynucleotides and/or polypeptides in a eukaryote, particularly a mammal, and especially a human, will provide a diagnostic method for diagnosis of disease, staging of disease or response of an infectious
25 organism to drugs. Eukaryotes, particularly mammals, and especially humans, particularly those infected or suspected to be infected with an organism comprising a BASB232 genes or proteins, may be detected at the nucleic acid or amino acid level by a variety of well known techniques as well as by methods provided herein.

Polypeptides and polynucleotides for prognosis, diagnosis or other analysis may be obtained from a putatively infected and/or infected individual's bodily materials. Polynucleotides from any of these sources, particularly DNA or RNA, may be used directly for detection or may be amplified enzymatically by using PCR or any other amplification technique prior to analysis. RNA, particularly mRNA, cDNA and genomic DNA may also be used in the same ways. Using amplification, characterization of the species and strain of infectious or resident organism present in an individual, may be made by an analysis of the genotype of a selected polynucleotide of the organism. Deletions and insertions can be detected by a change in size of the amplified product in comparison to a genotype of a reference sequence selected from a related organism, preferably a different species of the same genus or a different strain of the same species. Point mutations can be identified by hybridizing amplified DNA to labeled BASB232 polynucleotide sequences. Perfectly or significantly matched sequences can be distinguished from imperfectly or more significantly mismatched duplexes by DNase or RNase digestion, for DNA or RNA respectively, or by detecting differences in melting temperatures or renaturation kinetics. Polynucleotide sequence differences may also be detected by alterations in the electrophoretic mobility of polynucleotide fragments in gels as compared to a reference sequence. This may be carried out with or without denaturing agents. Polynucleotide differences may also be detected by direct DNA or RNA sequencing. See, for example, Myers *et al.*, *Science*, 230: 1242 (1985). Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase, V1 and S1 protection assay or a chemical cleavage method. See, for example, Cotton *et al.*, *Proc. Natl. Acad. Sci., USA*, 85: 4397-4401 (1985).

In another embodiment, an array of oligonucleotides probes comprising BASB232 nucleotide sequences or fragments thereof can be constructed to conduct efficient screening of, for example, genetic mutations, serotype, taxonomic classification or identification. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability (see, for example, Chee *et al.*, *Science*, 274: 610 (1996)).

Thus in another aspect, the present invention relates to a diagnostic kit which comprises:

- (a) a polynucleotide of the present invention, preferably any of the nucleotide sequences of SEQ Group 1, or a fragment thereof ;
 - (b) a nucleotide sequence complementary to that of (a);
 - 5 (c) a polypeptide of the present invention, preferably any of the polypeptides of SEQ Group 2 or a fragment thereof; or
 - (d) an antibody to a polypeptide of the present invention, preferably to the polypeptides of SEQ Group 2 .
- 10 It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component. Such a kit will be of use in diagnosing a disease or susceptibility to a Disease, among others.

- This invention also relates to the use of polynucleotides of the present invention as
- 15 diagnostic reagents. Detection of a mutated form of a polynucleotide of the invention, preferably any sequences of SEQ Group 1 , which is associated with a disease or pathogenicity will provide a diagnostic tool that can add to, or define, a diagnosis of a disease, a prognosis of a course of disease, a determination of a stage of disease, or a susceptibility to a disease, which results from under-expression, over-expression or altered
- 20 expression of the polynucleotide. Organisms, particularly infectious organisms, carrying mutations in such polynucleotide may be detected at the polynucleotide level by a variety of techniques, such as those described elsewhere herein.

- Cells from an organism carrying mutations or polymorphisms (allelic variations) in a
- 25 polynucleotide and/or polypeptide of the invention may also be detected at the polynucleotide or polypeptide level by a variety of techniques, to allow for serotyping, for example. For example, RT-PCR can be used to detect mutations in the RNA. It is particularly preferred to use RT-PCR in conjunction with automated detection systems, such as, for example, GeneScan. RNA, cDNA or genomic DNA may also be used for the same

purpose, PCR. As an example, PCR primers complementary to a polynucleotide encoding BASB232 polypeptide can be used to identify and analyze mutations.

5 The invention further provides primers with 1, 2, 3 or 4 nucleotides removed from the 5' and/or the 3' end. These primers may be used for, among other things, amplifying BASB232 DNA and/or RNA isolated from a sample derived from an individual, such as a bodily material. The primers may be used to amplify a polynucleotide isolated from an infected individual, such that the polynucleotide may then be subject to various techniques for elucidation of the polynucleotide sequence. In this way, mutations in the polynucleotide
10 sequence may be detected and used to diagnose and/or prognose the infection or its stage or course, or to serotype and/or classify the infectious agent.

The invention further provides a process for diagnosing, disease, preferably bacterial infections, more preferably infections caused by *B. pertussis*, comprising determining from
15 a sample derived from an individual, such as a bodily material, an increased level of expression of polynucleotide having a sequence of any of the sequences of SEQ Group 1. Increased or decreased expression of a BASB232 polynucleotides can be measured using any one of the methods well known in the art for the quantitation of polynucleotides, such as, for example, amplification, PCR, RT-PCR, RNase protection, Northern blotting,
20 spectrometry and other hybridization methods.

In addition, a diagnostic assay in accordance with the invention for detecting over-expression of BASB232 polypeptides compared to normal control tissue samples may be used to detect the presence of an infection, for example. Assay techniques that can be used
25 to determine levels of BASB232 polypeptides, in a sample derived from a host, such as a bodily material, are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis, antibody sandwich assays, antibody detection and ELISA assays.

The polynucleotides of the invention may be used as components of polynucleotide arrays, preferably high density arrays or grids. These high density arrays are particularly useful for diagnostic and prognostic purposes. For example, a set of spots each comprising a different gene, and further comprising a polynucleotide or polynucleotides of the invention, may be used for probing, such as using hybridization or nucleic acid amplification, using a probes obtained or derived from a bodily sample, to determine the presence of a particular polynucleotide sequence or related sequence in an individual. Such a presence may indicate the presence of a pathogen, particularly *B. pertussis*, and may be useful in diagnosing and/or prognosing disease or a course of disease. A grid comprising a number of variants of any polynucleotide sequences of SEQ Group 1 are preferred. Also preferred is a comprising a number of variants of a polynucleotide sequence encoding any polypeptide sequences of SEQ Group 2 .

Antibodies

The polypeptides and polynucleotides of the invention or variants thereof, or cells expressing the same can be used as immunogens to produce antibodies immunospecific for such polypeptides or polynucleotides respectively. Alternatively, mimotopes, particularly peptide mimotopes, of epitopes within the polypeptide sequence may also be used as immunogens to produce antibodies immunospecific for the polypeptide of the invention. The term "immunospecific" means that the antibodies have substantially greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art.

In certain preferred embodiments of the invention there are provided antibodies against BASB232 polypeptides or polynucleotides.

Antibodies generated against the polypeptides or polynucleotides of the invention can be obtained by administering the polypeptides and/or polynucleotides of the invention, or epitope-bearing fragments of either or both, analogues of either or both, or cells expressing either or both, to an animal, preferably a nonhuman, using routine protocols. For

preparation of monoclonal antibodies, any technique known in the art that provides antibodies produced by continuous cell line cultures can be used. Examples include various techniques, such as those in Kohler, G. and Milstein, C., *Nature* 256: 495-497 (1975); Kozbor *et al.*, *Immunology Today* 4: 72 (1983); Cole *et al.*, pg. 77-96 in *MONOCLONAL*
5 *ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc. (1985).

Techniques for the production of single chain antibodies (U.S. Patent No. 4,946,778) can be adapted to produce single chain antibodies to polypeptides or polynucleotides of this invention. Also, transgenic mice, or other organisms or animals, such as other mammals,
10 may be used to express humanized antibodies immunospecific to the polypeptides or polynucleotides of the invention.

Alternatively, phage display technology may be utilized to select antibody genes with binding activities towards a polypeptide of the invention either from repertoires of PCR
15 amplified v-genes of lymphocytes from humans screened for possessing anti-BASB232 or from naive libraries (McCafferty, *et al.*, (1990), *Nature* 348, 552-554; Marks, *et al.*, (1992) *Biotechnology* 10, 779-783). The affinity of these antibodies can also be improved by, for example, chain shuffling (Clackson *et al.*, (1991) *Nature* 352: 628).

20 The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptides or polynucleotides of the invention to purify the polypeptides or polynucleotides by, for example, affinity chromatography.

Thus, among others, antibodies against BASB232 polypeptides or BASB232
25 polynucleotides may be employed to treat infections, particularly bacterial infections.

Polypeptide variants include antigenically, epitopically or immunologically equivalent variants form a particular aspect of this invention.

Preferably, the antibody or variant thereof is modified to make it less immunogenic in the individual. For example, if the individual is human the antibody may most preferably be "humanized," where the complementarity determining region or regions of the hybridoma-derived antibody has been transplanted into a human monoclonal antibody, for example as described in Jones *et al.* (1986), *Nature* 321, 522-525 or Tempest *et al.*, (1991) *Biotechnology* 9, 266-273.

Antagonists and Agonists - Assays and Molecules

Polypeptides and polynucleotides of the invention may also be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics. See, *e.g.*, Coligan *et al.*, *Current Protocols in Immunology* 1(2): Chapter 5 (1991).

The screening methods may simply measure the binding of a candidate compound to the polypeptide or polynucleotide, or to cells or membranes bearing the polypeptide or polynucleotide, or a fusion protein of the polypeptide by means of a label directly or indirectly associated with the candidate compound. Alternatively, the screening method may involve competition with a labeled competitor. Further, these screening methods may test whether the candidate compound results in a signal generated by activation or inhibition of the polypeptide or polynucleotide, using detection systems appropriate to the cells comprising the polypeptide or polynucleotide. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed. Constitutively active polypeptide and/or constitutively expressed polypeptides and polynucleotides may be employed in screening methods for inverse agonists or inhibitors, in the absence of an agonist or inhibitor, by testing whether the candidate compound results in inhibition of activation of the polypeptide or polynucleotide, as the case may be. Further, the screening methods may simply comprise the steps of mixing a candidate compound with a solution

containing a polypeptide or polynucleotide of the present invention, to form a mixture, measuring BASB232 polypeptides and/or polynucleotides activity in the mixture, and comparing the BASB232 polypeptides and/or polynucleotides activity of the mixture to a standard. Fusion proteins, such as those made from Fc portion and BASB232

5 polypeptides, as hereinbefore described, can also be used for high-throughput screening assays to identify antagonists of the polypeptide of the present invention, as well as of phylogenetically and and/or functionally related polypeptides (see D. Bennett *et al.*, J Mol Recognition, 8:52-58 (1995); and K. Johanson *et al.*, J Biol Chem, 270(16):9459-9471 (1995)).

10

The polynucleotides, polypeptides and antibodies that bind to and/or interact with a polypeptide of the present invention may also be used to configure screening methods for detecting the effect of added compounds on the production of mRNA and/or polypeptide in cells. For example, an ELISA assay may be constructed for measuring secreted or cell
15 associated levels of polypeptide using monoclonal and polyclonal antibodies by standard methods known in the art. This can be used to discover agents which may inhibit or enhance the production of polypeptide (also called antagonist or agonist, respectively) from suitably manipulated cells or tissues.

20 The invention also provides a method of screening compounds to identify those which enhance (agonist) or block (antagonist) the action of BASB232 polypeptide or polynucleotides, particularly those compounds that are bacteriostatic and/or bactericidal. The method of screening may involve high-throughput techniques. For example, to screen for agonists or antagonists, a synthetic reaction mix, a cellular compartment, such as a
25 membrane, cell envelope or cell wall, or a preparation of any thereof, comprising BASB232 polypeptide and a labeled substrate or ligand of such polypeptides is incubated in the absence or the presence of a candidate molecule that may be a BASB232 agonist or antagonist. The ability of the candidate molecule to agonize or antagonize the BASB232 polypeptide is reflected in decreased binding of the labeled ligand or decreased production
30 of product from such substrate. Molecules that bind gratuitously, *i.e.*, without inducing the

effects of BASB232 polypeptide are most likely to be good antagonists. Molecules that bind well and, as the case may be, increase the rate of product production from substrate, increase signal transduction, or increase chemical channel activity are agonists. Detection of the rate or level of, as the case may be, production of product from substrate, signal transduction, or chemical channel activity may be enhanced by using a reporter system.

Reporter systems that may be useful in this regard include but are not limited to colorimetric, labeled substrate converted into product, a reporter gene that is responsive to changes in BASB232 polynucleotide or polypeptide activity, and binding assays known in the art.

Another example of an assay for BASB232 agonists is a competitive assay that combines BASB232 and a potential agonist with BASB232 binding molecules, recombinant BASB232 binding molecules, natural substrates or ligands, or substrate or ligand mimetics, under appropriate conditions for a competitive inhibition assay. BASB232 can be labeled, such as by radioactivity or a colorimetric compound, such that the number of BASB232 molecules bound to a binding molecule or converted to product can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include, among others, small organic molecules, peptides, polypeptides and antibodies that bind to a polynucleotide and/or polypeptide of the invention and thereby inhibit or extinguish its activity or expression. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as a binding molecule, without inducing BASB232 induced activities, thereby preventing the action or expression of BASB232 polypeptides and/or polynucleotides by excluding BASB232 polypeptides and/or polynucleotides from binding.

Potential antagonists include a small molecule that binds to and occupies the binding site of the polypeptide thereby preventing binding to cellular binding molecules, such that normal biological activity is prevented. Examples of small molecules include but are not limited to

small organic molecules, peptides or peptide-like molecules. Other potential antagonists include antisense molecules (see Okano, *J. Neurochem.* 56: 560 (1991);

OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION, CRC Press, Boca Raton, FL (1988), for a description of these molecules). Preferred

5 potential antagonists include compounds related to and variants of BASB232.

In a further aspect, the present invention relates to genetically engineered soluble fusion proteins comprising a polypeptide of the present invention, or a fragment thereof, and various portions of the constant regions of heavy or light chains of immunoglobulins of
10 various subclasses (IgG, IgM, IgA, IgE). Preferred as an immunoglobulin is the constant part of the heavy chain of human IgG, particularly IgG1, where fusion takes place at the hinge region. In a particular embodiment, the Fc part can be removed simply by incorporation of a cleavage sequence which can be cleaved with blood clotting factor Xa. Furthermore, this invention relates to processes for the preparation of these fusion
15 proteins by genetic engineering, and to the use thereof for drug screening, diagnosis and therapy. A further aspect of the invention also relates to polynucleotides encoding such fusion proteins. Examples of fusion protein technology can be found in International Patent Application Nos. WO94/29458 and WO94/22914.

20 Each of the polynucleotide sequences provided herein may be used in the discovery and development of antibacterial compounds. The encoded protein, upon expression, can be used as a target for the screening of antibacterial drugs. Additionally, the polynucleotide sequences encoding the amino terminal regions of the encoded protein or Shine-Delgarno or other translation facilitating sequences of the respective mRNA can be used to
25 construct antisense sequences to control the expression of the coding sequence of interest.

The invention also provides the use of the polypeptide, polynucleotide, agonist or antagonist of the invention to interfere with the initial physical interaction between a pathogen or pathogens and a eukaryotic, preferably mammalian, host responsible for
30 sequelae of infection. In particular, the molecules of the invention may be used: in the

prevention of adhesion of bacteria, in particular gram positive and/or gram negative bacteria, to eukaryotic, preferably mammalian, extracellular matrix proteins on in-dwelling devices or to extracellular matrix proteins in wounds; to block bacterial adhesion between eukaryotic, preferably mammalian, extracellular matrix proteins and bacterial BASB232 proteins that mediate tissue damage and/or; to block the normal progression of pathogenesis in infections initiated other than by the implantation of in-dwelling devices or by other surgical techniques.

In accordance with yet another aspect of the invention, there are provided BASB232 agonists and antagonists, preferably bacteristatic or bactericidal agonists and antagonists.

The antagonists and agonists of the invention may be employed, for instance, to prevent, inhibit and/or treat diseases.

In a further aspect, the present invention relates to mimotopes of the polypeptide of the invention. A mimotope is a peptide sequence, sufficiently similar to the native peptide (sequentially or structurally), which is capable of being recognised by antibodies which recognise the native peptide; or is capable of raising antibodies which recognise the native peptide when coupled to a suitable carrier.

Peptide mimotopes may be designed for a particular purpose by addition, deletion or substitution of elected amino acids. Thus, the peptides may be modified for the purposes of ease of conjugation to a protein carrier. For example, it may be desirable for some chemical conjugation methods to include a terminal cysteine. In addition it may be desirable for peptides conjugated to a protein carrier to include a hydrophobic terminus distal from the conjugated terminus of the peptide, such that the free unconjugated end of the peptide remains associated with the surface of the carrier protein. Thereby presenting the peptide in a conformation which most closely resembles that of the peptide as found in the context of the whole native molecule. For example, the peptides may be altered to have an N-terminal cysteine and a C-terminal hydrophobic amidated

tail. Alternatively, the addition or substitution of a D-stereoisomer form of one or more of the amino acids may be performed to create a beneficial derivative, for example to enhance stability of the peptide.

- 5 Alternatively, peptide mimotopes may be identified using antibodies which are capable themselves of binding to the polypeptides of the present invention using techniques such as phage display technology (EP 0 552 267 B1). This technique, generates a large number of peptide sequences which mimic the structure of the native peptides and are, therefore, capable of binding to anti-native peptide antibodies, but may not necessarily themselves
10 share significant sequence homology to the native polypeptide.

Vaccines

- Another aspect of the invention relates to a method for inducing an immunological response in an individual, particularly a mammal, preferably humans, which comprises
15 inoculating the individual with BASB232 polynucleotide and/or polypeptide, or a fragment or variant thereof, or a combination thereof as described above, adequate to produce antibody and/ or T cell immune response to protect said individual from infection, particularly bacterial infection and most particularly *B. pertussis* (and preferably *B. parapertussis*) infection. Also provided are methods whereby such immunological
20 response slows bacterial replication. Yet another aspect of the invention relates to a method of inducing immunological response in an individual which comprises delivering to such individual a nucleic acid vector, sequence or ribozyme to direct expression of BASB232 polynucleotide and/or polypeptide, or a fragment or a variant thereof, for expressing BASB232 polynucleotide and/or polypeptide, or a fragment or a variant
25 thereof, or a combination thereof as described above, *in vivo* in order to induce an immunological response, such as, to produce antibody and/ or T cell immune response, including, for example, cytokine-producing T cells or cytotoxic T cells, to protect said individual, preferably a human, from disease, whether that disease is already established within the individual or not. One example of administering the gene is by accelerating it
30 into the desired cells as a coating on particles or otherwise. Such nucleic acid vector may

comprise DNA, RNA, a ribozyme, a modified nucleic acid, a DNA/RNA hybrid, a DNA-protein complex or an RNA-protein complex.

5 A further aspect of the invention relates to an immunological composition that when introduced into an individual, preferably a human, capable of having induced within it an immunological response, induces an immunological response in such individual to a BASB232 polynucleotide and/or polypeptide encoded therefrom, or a combination thereof as described above, wherein the composition comprises a recombinant BASB232 polynucleotide and/or polypeptide encoded therefrom and/or comprises DNA and/or RNA
10 which encodes and expresses an antigen of said BASB232 polynucleotide, polypeptide encoded therefrom, or other polypeptide of the invention. The immunological response may be used therapeutically or prophylactically and may take the form of antibody immunity and/or cellular immunity, such as cellular immunity arising from CTL or CD4+ T cells.

15 A BASB232 polypeptide or a fragment thereof may be fused with co-protein or chemical moiety which may or may not by itself produce antibodies, but which is capable of stabilizing the first protein and producing a fused or modified protein which will have antigenic and/or immunogenic properties, and preferably protective properties. Thus
20 fused recombinant protein, preferably further comprises an antigenic co-protein, such as lipoprotein D from *Haemophilus influenzae*, Glutathione-S-transferase (GST) or beta-galactosidase, or any other relatively large co-protein which solubilizes the protein and facilitates production and purification thereof. Moreover, the co-protein may act as an adjuvant in the sense of providing a generalized stimulation of the immune system of the
25 organism receiving the protein. The co-protein may be attached to either the amino- or carboxy-terminus of the first protein.

In a vaccine composition according to the invention, a BASB232 polypeptide and/or polynucleotide, or a fragment, or a mimotope, or a variant thereof, or a combination thereof

as described above, may be present in a vector, such as the live recombinant vectors described above for example live bacterial vectors.

Also suitable are non-live vectors for the BASB232 polypeptide, or a combination thereof as described above, for example bacterial outer-membrane vesicles or "blebs". OM blebs are derived from the outer membrane of the two-layer membrane of Gram-negative bacteria and have been documented in many Gram-negative bacteria (Zhou, L *et al.* 1998. *FEMS Microbiol. Lett.* 163:223-228) including *C. trachomatis* and *C. psittaci*. A non-exhaustive list of bacterial pathogens reported to produce blebs also includes: *Bordetella pertussis*, *Borrelia burgdorferi*, *Brucella melitensis*, *Brucella ovis*, *Escherichia coli*, *Haemophilus influenzae*, *Legionella pneumophila*, *Moraxella catarrhalis*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Pseudomonas aeruginosa* and *Yersinia enterocolitica*.

Blebs have the advantage of providing outer-membrane proteins in their native conformation and are thus particularly useful for vaccines. Blebs can also be improved for vaccine use by engineering the bacterium so as to modify the expression of one or more molecules at the outer membrane. Thus for example the expression of a desired immunogenic protein at the outer membrane, such as the BASB232 polypeptide, can be introduced or upregulated (e.g. by altering the promoter). Instead or in addition, the expression of outer-membrane molecules which are either not relevant (e.g. unprotective antigens or immunodominant but variable proteins) or detrimental (e.g. toxic molecules such as LPS, or potential inducers of an autoimmune response) can be downregulated. These approaches are discussed in more detail below.

The non-coding flanking regions of the BASB232 genes contain regulatory elements important in the expression of the gene. This regulation takes place both at the transcriptional and translational level. The sequence of these regions, either upstream or downstream of the open reading frame of the gene, can be obtained by DNA sequencing.

This sequence information allows the determination of potential regulatory motifs such as

the different promoter elements, terminator sequences, inducible sequence elements, repressors, elements responsible for phase variation, the shine-dalgarno sequence, regions with potential secondary structure involved in regulation, as well as other types of regulatory motifs or sequences. This sequence is a further aspect of the invention.

5

This sequence information allows the modulation of the natural expression of the BASB232 genes. The upregulation of the gene expression may be accomplished by altering the promoter, the shine-dalgarno sequence, potential repressor or operator elements, or any other elements involved. Likewise, downregulation of expression can be achieved by similar types of modification. Alternatively, by changing phase variation sequences, the expression of the gene can be put under phase variation control, or it may be uncoupled from this regulation. In another approach, the expression of the gene can be put under the control of one or more inducible elements allowing regulated expression. Examples of such regulation include, but are not limited to, induction by temperature shift, addition of inductor substrates like selected carbohydrates or their derivatives, trace elements, vitamins, co-factors, metal ions, etc.

10

15

20

25

30

Such modifications as described above can be introduced by several different means. The modification of sequences involved in gene expression can be carried out *in vivo* by random mutagenesis followed by selection for the desired phenotype. Another approach consists in isolating the region of interest and modifying it by random mutagenesis, or site-directed replacement, insertion or deletion mutagenesis. The modified region can then be reintroduced into the bacterial genome by homologous recombination, and the effect on gene expression can be assessed. In another approach, the sequence knowledge of the region of interest can be used to replace or delete all or part of the natural regulatory sequences. In this case, the regulatory region targeted is isolated and modified so as to contain the regulatory elements from another gene, a combination of regulatory elements from different genes, a synthetic regulatory region, or any other regulatory region, or to delete selected parts of the wild-type regulatory sequences. These modified sequences can then be reintroduced into the bacterium via homologous recombination into the genome.

A non-exhaustive list of preferred promoters that could be used for up-regulation of gene expression includes the promoters *porA*, *porB*, *lbpB*, *tbpB*, *p110*, *lst*, *hpuAB* from *N. meningitidis* or *N. gonorrhoeae*; *ompCD*, *copB*, *lbpB*, *ompE*, *UspA1*; *UspA2*; *TbpB* from *M. Catarrhalis*; *p1*, *p2*, *p4*, *p5*, *p6*, *lpD*, *tbpB*, *D15*, *Hia*, *Hmw1*, *Hmw2* from *H.*

5 *influenzae*.

In one example, the expression of the gene can be modulated by exchanging its promoter with a stronger promoter (through isolating the upstream sequence of the gene, in vitro modification of this sequence, and reintroduction into the genome by homologous recombination). Upregulated expression can be obtained in both the bacterium as well as
10 in the outer membrane vesicles shed (or made) from the bacterium.

In other examples, the described approaches can be used to generate recombinant bacterial strains with improved characteristics for vaccine applications. These can be, but are not limited to, attenuated strains, strains with increased expression of selected antigens,
15 strains with knock-outs (or decreased expression) of genes interfering with the immune response, strains with modulated expression of immunodominant proteins, strains with modulated shedding of outer-membrane vesicles.

Thus, also provided by the invention is a modified upstream region of the BASB232
20 genes, which modified upstream region contains a heterologous regulatory element which alters the expression level of the BASB232 proteins located at the outer membrane. The upstream region according to this aspect of the invention includes the sequence upstream of the BASB232 genes. The upstream region starts immediately upstream of the BASB232 genes and continues usually to a position no more than about 1000 bp upstream of the gene
25 from the ATG start codon. In the case of a gene located in a polycistronic sequence (operon) the upstream region can start immediately preceding the gene of interest, or preceding the first gene in the operon. Preferably, a modified upstream region according to this aspect of the invention contains a heterologous promotor at a position between 500 and 700 bp upstream of the ATG.

30

Thus, the invention provides BASB232 polypeptides, or a combination thereof as described above, in a modified bacterial bleb. The invention further provides modified host cells capable of producing the non-live membrane-based bleb vectors. The invention further provides nucleic acid vectors comprising the BASB232 genes having a modified upstream
5 region containing a heterologous regulatory element.

Further provided by the invention are processes to prepare the host cells and bacterial blebs according to the invention.

10 Also provided by this invention are compositions, particularly vaccine compositions, and methods comprising the polypeptides and/or polynucleotides of the invention and immunostimulatory DNA sequences, such as those described in Sato, Y. *et al.* Science 273: 352 (1996).

15 Also, provided by this invention are methods using the described polynucleotide or particular fragments thereof, which have been shown to encode non-variable regions of bacterial cell surface proteins, in polynucleotide constructs used in such genetic immunization experiments in animal models of infection with *B. pertussis*. Such experiments will be particularly useful for identifying protein epitopes able to provoke a
20 prophylactic or therapeutic immune response. It is believed that this approach will allow for the subsequent preparation of monoclonal antibodies of particular value, derived from the requisite organ of the animal successfully resisting or clearing infection, for the development of prophylactic agents or therapeutic treatments of bacterial infection, particularly *B. pertussis* infection, in mammals, particularly humans.

25

Immunogenic compositions

It is advantageous for immunogenic compositions of the invention to comprise immunogenic, preferably immunologically effective, amounts of additional antigens to
30 elicit immunity to other pathogens, preferably viruses and/or bacteria. Such additional

antigens include diphtheria toxoid, tetanus toxoid, hepatitis B surface antigen, injectable polio vaccine, *Haemophilus influenzae* type b PRP, fimbria 2 and fimbria 3.

Preferred immunogenic compositions of the invention are formulated with 1, 2, 3 or preferably all 4 of the following meningococcal capsular polysaccharides or oligosaccharides : A, C, Y or W, which may be plain or conjugated to a protein carrier. Such a vaccine containing proteins from *N. meningitidis* serogroup B may be advantageously combine a global meningococcus vaccine with a Bordetella vaccine.

10 In a further preferred embodiment, the immunogenic compositions of the invention are formulated with a conjugated or unconjugated *H. influenzae* b capsular polysaccharide or oligosaccharide, and one or more plain or conjugated pneumococcal capsular polysaccharides or oligosaccharides. Optionally, the vaccine may also comprise one or more protein antigens that can protect a host against *Streptococcus pneumoniae* 15 infection. Such a vaccine may be advantageously used as a Bordetella/H. influenzae/streptococcus pneumonia vaccine.

In a still further preferred embodiment, the immunogenic composition of the invention is formulated with capsular polysaccharides or oligosaccharides derived from one or 20 more of *Neisseria meningitidis*, *Haemophilus influenzae*, *Streptococcus pneumoniae*, Group A Streptococci, Group B Streptococci, *Staphylococcus aureus* or *Staphylococcus epidermidis*. In a preferred embodiment, the immunogenic composition would comprise capsular polysaccharides or oligosaccharides derived from one or more of serogroups A, C, W and Y of *Neisseria meningitidis*. A further preferred embodiment would 25 comprise capsular polysaccharides or oligosaccharides derived from *Streptococcus pneumoniae*. The pneumococcal capsular polysaccharide or oligosaccharide antigens are preferably selected from serotypes 1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19A, 19F, 20, 22F, 23F and 33F (most preferably from serotypes 1, 3, 4, 5, 6B, 7F, 9V, 14, 18C, 19F and 23F). A further preferred embodiment would contain 30 the PRP capsular polysaccharides or oligosaccharides of *Haemophilus influenzae*. A

- further preferred embodiment would contain the Type 5, Type 8 or 336 capsular polysaccharides or oligosaccharides of *Staphylococcus aureus*. A further preferred embodiment would contain the Type I, Type II or Type III capsular polysaccharides of *Staphylococcus epidermidis*. A further preferred embodiment would contain the Type Ia, Type Ic, Type II or Type III capsular polysaccharides or oligosaccharides of Group B streptococcus. A further preferred embodiment would contain the capsular polysaccharides or oligosaccharides of Group A streptococcus, preferably further comprising at least one M protein and more preferably multiple types of M protein.
- 10 Capsular polysaccharides or oligosaccharides included in pharmaceutical compositions of the invention may be unconjugated or conjugated to a carrier protein such as tetanus toxoid, tetanus toxoid fragment C, diphtheria toxoid, CRM197, pneumolysin, Protein D (US6342224).
- 15 The polysaccharide conjugate may be prepared by any known coupling technique. For example the polysaccharide can be coupled via a thioether linkage. This conjugation method relies on activation of the polysaccharide with 1-cyano-4-dimethylamino pyridinium tetrafluoroborate (CDAP) to form a cyanate ester. The activated polysaccharide may thus be coupled directly or via a spacer group to an amino group on the carrier protein. Preferably, the cyanate ester is coupled with hexane diamine and the amino-derivatised polysaccharide is conjugated to the carrier protein using heterologation chemistry involving the formation of the thioether linkage. Such conjugates are described in PCT published application WO93/15760 Uniformed Services University.
- 20 The conjugates can also be prepared by direct reductive amination methods as described in US 4365170 (Jennings) and US 4673574 (Anderson). Other methods are described in EP-0-161-188, EP-208375 and EP-0-477508.
- 25 A further method involves the coupling of a cyanogen bromide activated polysaccharide derivatised with adipic acid hydrazide (ADH) to the protein carrier by Carbodiimide condensation (Chu C. et al Infect. Immunity, 1983 245 256).
- 30

The immunogenic compositions of the invention may also comprise proteins from other pathogens. Preferred pneumococcal proteins antigens are those pneumococcal proteins which are exposed on the outer surface of the pneumococcus (capable of being
 5 recognised by a host's immune system during at least part of the life cycle of the pneumococcus), or are proteins which are secreted or released by the pneumococcus.

Most preferably, the protein is a toxin, adhesin, 2-component signal transducer, or lipoprotein of *Streptococcus pneumoniae*, or fragments thereof. Particularly preferred proteins include, but are not limited to: pneumolysin (preferably detoxified by chemical

10 treatment or mutation) [Mitchell *et al.* Nucleic Acids Res. 1990 Jul 11; 18(13): 4010 "Comparison of pneumolysin genes and proteins from *Streptococcus pneumoniae* types 1 and 2.", Mitchell *et al.* Biochim Biophys Acta 1989 Jan 23; 1007(1): 67-72

"Expression of the pneumolysin gene in *Escherichia coli*: rapid purification and biological properties.", WO 96/05859 (A. Cyanamid), WO 90/06951 (Paton et al), WO

15 99/03884 (NAVA)]; PspA and transmembrane deletion variants thereof (US 5804193 - Briles *et al.*); PspC and transmembrane deletion variants thereof (WO 97/09994 - Briles et al); PsaA and transmembrane deletion variants thereof (Berry & Paton, Infect Immun 1996 Dec;64(12):5255-62 "Sequence heterogeneity of PsaA, a 37-kilodalton putative adhesin essential for virulence of *Streptococcus pneumoniae*"); pneumococcal choline
 20 binding proteins and transmembrane deletion variants thereof; CbpA and transmembrane deletion variants thereof (WO 97/41151; WO 99/51266);

Glyceraldehyde-3-phosphate – dehydrogenase (Infect. Immun. 1996 64:3544); HSP70 (WO 96/40928); PcpA (Sanchez-Beato et al. *FEMS Microbiol Lett* 1998, 164:207-14);

M like protein, (EP 0837130) and adhesin 18627, (EP 0834568). Further preferred

25 pneumococcal protein antigens are those disclosed in WO 98/18931, particularly those selected in WO 98/18930 and PCT/US99/30390.

Preferred proteins for inclusion in the immunogenic composition of the invention include adhesins, autotransporter proteins, iron acquisition proteins and toxins from *N.*

30 *meningitidis* serotype B.

- Adhesins include FhaB (WO98/02547), NadA (J. Exp.Med (2002) 195:1445; NMB 1994), Hsf also known as NhhA (NMB 0992) (WO99/31132), Hap (NMB 1985)(WO99/55873), NspA (WO96/29412), MafA (NMB 0652) and MafB (NMB 0643) (Annu Rev Cell Dev Biol. 16; 423-457 (2000); Nature Biotech 20; 914-921 (2002)) , Omp26 (NMB 0181), NMB 0315, NMB 0995, NMB 1119 and PilC (Mol. Microbiol.1997, 23; 879-892). These are proteins that are involved in the binding of Neisseria to the surface of host cells.
- 10 Autotransporter proteins typically are made up of a signal sequence, a passenger domain and an anchoring domain for attachment to the outer membrane. Examples of autotransporter proteins include Hsf (WO99/31132) (NMB 0992), HMW, Hia (van Ulsen et al Immunol. Med. Microbiol. 2001 32; 53-64), Hap (NMB 1985) (WO99/55873; van Ulsen et al Immunol. Med. Microbiol. 2001 32; 53-64), UspA, 15 UspA2, NadA (NMB 1994) (Comanducci et al J. Exp. Med. 2002 195; 1445-1454), AspA (Infection and Immunity 2002, 70(8); 4447-4461; NMB 1029), Aida-1 like protein, SSh-2 and Tsh. The passenger domain of an autotransporter protein is a preferred fragment for incorporation into the immunogenic composition of the invention.
- 20 Iron acquisition proteins include TbpA (NMB 0461) (WO92/03467, US5912336, WO93/06861 and EP586266), TbpB (NMB 0460) (WO93/06861 and EP586266), LbpA (NMB 1540) (Med Microbiol (1999) 32:1117), LbpB (NMB 1541)(WO/99/09176), HpuA (U73112.2) (Mol Microbiol. 1997, 23; 737-749), HpuB (NC_003116.1) (Mol 25 Microbiol. 1997, 23; 737-749), P2086 also known as XthA (NMB 0399) (13th International Pathogenic Neisseria Conference 2002), FbpA (NMB 0634), FbpB, BfrA (NMB 1207), BfrB (NMB 1206), Lipo28 also known as GNA2132 (NMB 2132), Sibp (NMB 1882), HmbR, HemH, Bcp (NMB 0750), Iron (III) ABC transporter-permease protein (Tettelin et al Science 287; 1809-1815 2000), Iron (III) ABC transporter – 30 periplasmic (Tettelin et al Science 287; 1809-1815 2000), TonB-dependent receptor

(NMB 0964 and NMB 0293)(Tettelin et al Science 287; 1809-1815 2000) and transferrin binding protein related protein (Tettelin et al Science 287; 1809-1815 2000).

Toxins include FrpA (NMB 0585; NMB 1405), FrpA/C (see below for definition),
 5 FrpC (NMB 1415; NMB 1405) (WO92/01460), NM-ADPRT (NMB 1343) (13th
 International Pathogenic Neisseria Conference 2002 Masignani et al p135), VapD
 (NMB 1753), lipopolysaccharide (LPS; also called lipooligosaccharide or LOS)
 immunotype L2 and LPS immunotype L3. FrpA and FrpC contain a region which is
 conserved between these two proteins and a preferred fragment of the proteins would be
 10 a polypeptide containing this conserved fragment, preferably comprising amino acids
 227-1004 of the sequence of FrpA/C.

The meningococcal proteins included in the immunogenic composition of the invention
 may be present as a subunit composition in which the purified protein or an
 15 immunogenic fragment of the protein is added to the immunogenic composition.
 Preferably, the protein may be added as part of an outer membrane vesicle preparation.

The immunogenic composition may also optionally comprise antigens providing
 protection against Diphtheria and/or tetanus infections. Typically, the antigens
 20 providing protection against Diphtheria and tetanus would be Diphtheria toxoid and
 tetanus toxoid. The toxoids may be chemically inactivated toxins or toxins inactivated
 by the introduction of point mutations.

It is advantageous to combine the immunogenic composition of the invention with
 25 antigens that confer immunity against one or more of Haemophilus influenzae b,
 hepatitis B and/or polio virus. Preferred pharmaceutical compositions of the invention
 will further comprise one or more , most preferably all three of PRP polysaccharide or
 oligosaccharide of Haemophilus influenzae b, hepatitis B surface antigen and/or
 injectable polio virus (IPV).

30

The immunogenic composition may also optionally comprise one or more antigens that can protect a host against RSV and/or one or more antigens that can protect a host against influenza virus.

5 Preferred influenza virus antigens include whole, live or inactivated virus, split influenza virus, grown in eggs or MDCK cells, or Vero cells or whole flu virosomes (as described by R. Gluck, Vaccine, 1992, 10, 915-920) or purified or recombinant proteins thereof, such as HA, NP, NA, or M proteins, or combinations thereof.

Preferred RSV (Respiratory Syncytial Virus) antigens include the F glycoprotein, the G glycoprotein, the HN protein, or derivatives thereof.

10

Preferred non-typeable *H. influenzae* protein antigens include Fimbrin protein (US 5766608) and fusions comprising peptides therefrom (eg LB1 Fusion) (US 5843464 - Ohio State Research Foundation), OMP26, P6, protein D, TbpA, TbpB, Hia, Hmw1, Hmw2, Hap, and D15.

15

It should be appreciated that immunogenic compositions of the invention may comprise one or more capsular polysaccharide or oligosaccharide from a single species of bacteria. Immunogenic compositions may also comprise capsular polysaccharides or oligosaccharide derived from one or more species of bacteria.

20

Vaccines

A further embodiment of the invention provides a vaccine formulation which comprises an immunogenic recombinant polypeptide and/or polynucleotide of the invention, or a combination thereof, together with a suitable carrier/excipient, such as a pharmaceutically acceptable carrier/excipient. Since the polypeptides and polynucleotides may be broken down in the stomach, each is preferably administered parenterally, including, for example, administration that is subcutaneous, intramuscular, intravenous, or intradermal.

25

Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostatic

30

compounds and solutes which render the formulation isotonic with the bodily fluid,

preferably the blood, of the individual; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile
5 liquid carrier immediately prior to use.

The vaccine formulation of the invention may also include adjuvant systems for enhancing the immunogenicity of the formulation. Preferably the adjuvant system raises preferentially a TH1 type of response.

10

An immune response may be broadly distinguished into two extreme categories, being a humoral or cell mediated immune responses (traditionally characterised by antibody and cellular effector mechanisms of protection respectively). These categories of response have been termed TH1-type responses (cell-mediated response), and TH2-type immune
15 responses (humoral response).

20

Extreme TH1-type immune responses may be characterised by the generation of antigen specific, haplotype restricted cytotoxic T lymphocytes, and natural killer cell responses. In mice TH1-type responses are often characterised by the generation of antibodies of the IgG2a subtype, whilst in the human these correspond to IgG1 type antibodies. TH2-type immune responses are characterised by the generation of a broad range of immunoglobulin isotypes including in mice IgG1, IgA, and IgM.

25

It can be considered that the driving force behind the development of these two types of immune responses are cytokines. High levels of TH1-type cytokines tend to favour the induction of cell mediated immune responses to the given antigen, whilst high levels of TH2-type cytokines tend to favour the induction of humoral immune responses to the antigen.

The distinction of TH1 and TH2-type immune responses is not absolute. In reality an individual will support an immune response which is described as being predominantly TH1 or predominantly TH2. However, it is often convenient to consider the families of cytokines in terms of that described in murine CD4 +ve T cell clones by Mosmann and
5 Coffman (*Mosmann, T.R. and Coffman, R.L. (1989) TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. Annual Review of Immunology, 7, p145-173*). Traditionally, TH1-type responses are associated with the production of the INF- γ and IL-2 cytokines by T-lymphocytes. Other cytokines often directly associated with the induction of TH1-type immune responses are not produced
10 by T-cells, such as IL-12. In contrast, TH2-type responses are associated with the secretion of IL-4, IL-5, IL-6 and IL-13.

It is known that certain vaccine adjuvants are particularly suited to the stimulation of either TH1 or TH2 - type cytokine responses. Traditionally the best indicators of the
15 TH1:TH2 balance of the immune response after a vaccination or infection includes direct measurement of the production of TH1 or TH2 cytokines by T lymphocytes *in vitro* after restimulation with antigen, and/or the measurement of the IgG1:IgG2a ratio of antigen specific antibody responses.

20 Thus, a TH1-type adjuvant is one which preferentially stimulates isolated T-cell populations to produce high levels of TH1-type cytokines when re-stimulated with antigen *in vitro*, and promotes development of both CD8+ cytotoxic T lymphocytes and antigen specific immunoglobulin responses associated with TH1-type isotype.

25 Adjuvants which are capable of preferential stimulation of the TH1 cell response are described in International Patent Application No. WO 94/00153 and WO 95/17209.

3 De-O-acylated monophosphoryl lipid A (3D-MPL) is one such adjuvant. This is known from GB 2220211 (Ribi). Chemically it is a mixture of 3 De-O-acylated
30 monophosphoryl lipid A with 4, 5 or 6 acylated chains and is manufactured by Ribi

Immunochem, Montana. A preferred form of 3 De-O-acylated monophosphoryl lipid A is disclosed in European Patent 0 689 454 B1 (SmithKline Beecham Biologicals SA).

Preferably, the particles of 3D-MPL are small enough to be sterile filtered through a 0.22micron membrane (European Patent number 0 689 454).

3D-MPL will be present in the range of 10µg - 100µg preferably 25-50µg per dose wherein the antigen will typically be present in a range 2-50µg per dose.

Another preferred adjuvant comprises QS21, an Hplc purified non-toxic fraction derived from the bark of *Quillaja Saponaria Molina*. Optionally this may be admixed with 3 De-O-acylated monophosphoryl lipid A (3D-MPL), optionally together with an carrier.

The method of production of QS21 is disclosed in US patent No. 5,057,540.

Non-reactogenic adjuvant formulations containing QS21 have been described previously (WO 96/33739). Such formulations comprising QS21 and cholesterol have been shown to be successful TH1 stimulating adjuvants when formulated together with an antigen.

Further adjuvants which are preferential stimulators of TH1 cell response include immunomodulatory oligonucleotides, for example unmethylated CpG sequences as disclosed in WO 96/02555.

Combinations of different TH1 stimulating adjuvants, such as those mentioned hereinabove, are also contemplated as providing an adjuvant which is a preferential stimulator of TH1 cell response. For example, QS21 can be formulated together with 3D-MPL. The ratio of QS21 : 3D-MPL will typically be in the order of 1 : 10 to 10 : 1; preferably 1:5 to 5 : 1 and often substantially 1 : 1. The preferred range for optimal synergy is 2.5 : 1 to 1 : 1 3D-MPL: QS21.

Preferably a carrier is also present in the vaccine composition according to the invention. The carrier may be an oil in water emulsion, or an aluminium salt, such as aluminium phosphate or aluminium hydroxide.

5

A preferred oil-in-water emulsion comprises a metabolisable oil, such as squalene, alpha tocopherol and Tween 80. In a particularly preferred aspect the antigens in the vaccine composition according to the invention are combined with QS21 and 3D-MPL in such an emulsion. Additionally the oil in water emulsion may contain span 85 and/or lecithin and/or tricaprylin.

10

Typically for human administration QS21 and 3D-MPL will be present in a vaccine in the range of 1µg - 200µg, such as 10-100µg, preferably 10µg - 50µg per dose.

Typically the oil in water will comprise from 2 to 10% squalene, from 2 to 10% alpha tocopherol and from 0.3 to 3% tween 80. Preferably the ratio of squalene: alpha tocopherol is equal to or less than 1 as this provides a more stable emulsion. Span 85 may also be present at a level of 1%. In some cases it may be advantageous that the vaccines of the present invention will further contain a stabiliser.

15

Non-toxic oil in water emulsions preferably contain a non-toxic oil, e.g. squalane or squalene, an emulsifier, e.g. Tween 80, in an aqueous carrier. The aqueous carrier may be, for example, phosphate buffered saline.

20

A particularly potent adjuvant formulation involving QS21, 3D-MPL and tocopherol in an oil in water emulsion is described in WO 95/17210.

25

While the invention has been described with reference to certain BASB232 polypeptides and polynucleotides, it is to be understood that this covers fragments of the naturally occurring polypeptides and polynucleotides, and similar polypeptides and polynucleotides

with additions, deletions or substitutions which do not substantially affect the immunogenic properties of the recombinant polypeptides or polynucleotides.

Compositions, kits and administration

- 5 In a further aspect of the invention there are provided compositions comprising a BASB232 polynucleotide and/or a BASB232 polypeptide for administration to a cell or to a multicellular organism.

- 10 The invention also relates to compositions comprising a polynucleotide and/or a polypeptides discussed herein or their agonists or antagonists. The polypeptides and polynucleotides of the invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to an individual. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of a polypeptide and/or
- 15 polynucleotide of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation should suit the mode of administration. The invention further relates to diagnostic and pharmaceutical packs and kits comprising one or more containers filled with 1, 2, 3, 4 or 5 of the ingredients of the aforementioned
- 20 compositions of the invention.

Polypeptides, polynucleotides and other compounds of the invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

- 25 The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes among others.

In therapy or as a prophylactic, the active agent may be administered to an individual as an injectable composition, for example as a sterile aqueous dispersion, preferably isotonic.

5 In a further aspect, the present invention provides for pharmaceutical compositions comprising a therapeutically effective amount of a polypeptide and/or polynucleotide, such as the soluble form of a polypeptide and/or polynucleotide of the present invention, agonist or antagonist peptide or small molecule compound, in combination with a pharmaceutically acceptable carrier or excipient. Such carriers include, but are not limited to, saline, buffered
10 saline, dextrose, water, glycerol, ethanol, and combinations thereof. The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention. Polypeptides, polynucleotides and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

15

The composition will be adapted to the route of administration, for instance by a systemic or an oral route. Preferred forms of systemic administration include injection, typically by intravenous injection. Other injection routes, such as subcutaneous, intramuscular, or intraperitoneal, can be used. Alternative means for systemic administration include
20 transmucosal and transdermal administration using penetrants such as bile salts or fusidic acids or other detergents. In addition, if a polypeptide or other compounds of the present invention can be formulated in an enteric or an encapsulated formulation, oral administration may also be possible. Administration of these compounds may also be topical and/or localized, in the form of salves, pastes, gels, solutions, powders and the like.

25

For administration to mammals, and particularly humans, it is expected that the daily dosage level of the active agent will be from 0.01 mg/kg to 10 mg/kg, typically around 1 mg/kg. The physician in any event will determine the actual dosage which will be most suitable for an individual and will vary with the age, weight and response of the particular
30 individual. The above dosages are exemplary of the average case. There can, of course,

be individual instances where higher or lower dosage ranges are merited, and such are within the scope of this invention.

5 The dosage range required depends on the choice of peptide, the route of administration, the nature of the formulation, the nature of the subject's condition, and the judgment of the attending practitioner. Suitable dosages, however, are in the range of 0.1-100 $\mu\text{g/kg}$ of subject.

10 A vaccine composition is conveniently in injectable form. Conventional adjuvants may be employed to enhance the immune response. A suitable unit dose for vaccination is 0.5-5 microgram/kg of antigen, and such dose is preferably administered 1-3 times and with an interval of 1-3 weeks. With the indicated dose range, no adverse toxicological effects will be observed with the compounds of the invention which would preclude their administration to suitable individuals.

15 Wide variations in the needed dosage, however, are to be expected in view of the variety of compounds available and the differing efficiencies of various routes of administration. For example, oral administration would be expected to require higher dosages than administration by intravenous injection. Variations in these dosage levels can be adjusted
20 using standard empirical routines for optimization, as is well understood in the art.

Sequence Databases, Sequences in a Tangible Medium, and Algorithms

25 Polynucleotide and polypeptide sequences form a valuable information resource with which to determine their 2- and 3-dimensional structures as well as to identify further sequences of similar homology. These approaches are most easily facilitated by storing the sequence in a computer readable medium and then using the stored data in a known macromolecular structure program or to search a sequence database using well known searching tools, such as the GCG program package.

Also provided by the invention are methods for the analysis of character sequences or strings, particularly genetic sequences or encoded protein sequences. Preferred methods of sequence analysis include, for example, methods of sequence homology analysis, such as identity and similarity analysis, DNA, RNA and protein structure analysis, sequence
5 assembly, cladistic analysis, sequence motif analysis, open reading frame determination, nucleic acid base calling, codon usage analysis, nucleic acid base trimming, and sequencing chromatogram peak analysis.

10 A computer based method is provided for performing homology identification. This method comprises the steps of: providing a first polynucleotide sequence comprising the sequence of a polynucleotide of the invention in a computer readable medium; and comparing said first polynucleotide sequence to at least one second polynucleotide or polypeptide sequence to identify homology.

15 A computer based method is also provided for performing homology identification, said method comprising the steps of: providing a first polypeptide sequence comprising the sequence of a polypeptide of the invention in a computer readable medium; and comparing said first polypeptide sequence to at least one second polynucleotide or polypeptide sequence to identify homology.

20

All publications and references, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference in their entirety as if each individual publication or reference were specifically and individually indicated to be incorporated by reference herein as being fully set forth. Any patent
25 application to which this application claims priority is also incorporated by reference herein in its entirety in the manner described above for publications and references.

DEFINITIONS

"Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as the case may be, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" can be readily calculated by known methods, including but not limited to those described in (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data*, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heine, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., *SIAM J. Applied Math.*, 48: 1073 (1988). Methods to determine identity are designed to give the largest match between the sequences tested. Moreover, methods to determine identity are codified in publicly available computer programs. Computer program methods to determine identity between two sequences include, but are not limited to, the GAP program in the GCG program package (Devereux, J., et al., *Nucleic Acids Research* 12(1): 387 (1984)), BLASTP, BLASTN (Altschul, S.F. et al., *J. Molec. Biol.* 215: 403-410 (1990), and FASTA(Pearson and Lipman *Proc. Natl. Acad. Sci. USA* 85: 2444-2448 (1988). The BLAST family of programs is publicly available from NCBI and other sources (*BLAST Manual*, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., *J. Mol. Biol.* 215: 403-410 (1990). The well known Smith Waterman algorithm may also be used to determine identity.

Parameters for polypeptide sequence comparison include the following:

Algorithm: Needleman and Wunsch, *J. Mol Biol.* 48: 443-453 (1970)

Comparison matrix: BLOSSUM62 from Henikoff and Henikoff,

Proc. Natl. Acad. Sci. USA. 89:10915-10919 (1992)

Gap Penalty: 8

Gap Length Penalty: 2

A program useful with these parameters is publicly available as the "gap" program from Genetics Computer Group, Madison WI. The aforementioned parameters are the default parameters for peptide comparisons (along with no penalty for end gaps).

Parameters for polynucleotide comparison include the following:

Algorithm: Needleman and Wunsch, J. Mol Biol. 48: 443-453 (1970)

Comparison matrix: matches = +10, mismatch = 0

10 Gap Penalty: 50

Gap Length Penalty: 3

Available as: The "gap" program from Genetics Computer Group, Madison WI. These are the default parameters for nucleic acid comparisons.

15 A preferred meaning for "identity" for polynucleotides and polypeptides, as the case may be, are provided in (1) and (2) below.

(1) Polynucleotide embodiments further include an isolated polynucleotide comprising a polynucleotide sequence having at least a 50, 60, 70, 80, 85, 90, 95, 97 or 100% identity to the reference sequence of SEQ ID NO:1, wherein said polynucleotide sequence may be identical to the reference sequence of SEQ ID NO:1 or may include up to a certain integer number of nucleotide alterations as compared to the reference sequence, wherein said alterations are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, or insertion, and wherein said alterations may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence, and wherein said number of nucleotide alterations is determined by multiplying the total number of nucleotides in SEQ ID NO:1

by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of nucleotides in SEQ ID NO:1, or:

$$n_n \leq x_n - (x_n \cdot y),$$

5

wherein n_n is the number of nucleotide alterations, x_n is the total number of nucleotides in SEQ ID NO:1, y is 0.50 for 50%, 0.60 for 60%, 0.70 for 70%, 0.80 for 80%, 0.85 for 85%, 0.90 for 90%, 0.95 for 95%, 0.97 for 97% or 1.00 for 100%, and \cdot is the symbol for the multiplication operator, and wherein any non-integer product of x_n and y is rounded
10 down to the nearest integer prior to subtracting it from x_n . Alterations of a polynucleotide sequence encoding the polypeptide of SEQ ID NO:2 may create nonsense, missense or frameshift mutations in this coding sequence and thereby alter the polypeptide encoded by the polynucleotide following such alterations.

15 By way of example, a polynucleotide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:1, that is it may be 100% identical, or it may include up to a certain integer number of nucleic acid alterations as compared to the reference sequence such that the percent identity is less than 100% identity. Such alterations are selected from the group consisting of at least one nucleic acid deletion,
20 substitution, including transition and transversion, or insertion, and wherein said alterations may occur at the 5' or 3' terminal positions of the reference polynucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleic acids in the reference sequence or in one or more contiguous groups within the reference sequence. The number of nucleic acid alterations for a given percent
25 identity is determined by multiplying the total number of nucleic acids in SEQ ID NO:1 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of nucleic acids in SEQ ID NO:1, or:

$$n_n \leq x_n - (x_n \cdot y),$$

wherein n_n is the number of nucleic acid alterations, x_n is the total number of nucleic acids in SEQ ID NO:1, y is, for instance 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc., \bullet is the symbol for the multiplication operator, and wherein any non-integer product of x_n and y is rounded down to the nearest integer prior to subtracting it from x_n .

(2) Polypeptide embodiments further include an isolated polypeptide comprising a polypeptide having at least a 50,60, 70, 80, 85, 90, 95, 97 or 100% identity to a polypeptide reference sequence of SEQ ID NO:2, wherein said polypeptide sequence may be identical to the reference sequence of SEQ ID NO:2 or may include up to a certain integer number of amino acid alterations as compared to the reference sequence, wherein said alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence, and wherein said number of amino acid alterations is determined by multiplying the total number of amino acids in SEQ ID NO:2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of amino acids in SEQ ID NO:2, or:

$$n_a \leq x_a - (x_a \bullet y),$$

wherein n_a is the number of amino acid alterations, x_a is the total number of amino acids in SEQ ID NO:2, y is 0.50 for 50%, 0.60 for 60%, 0.70 for 70%, 0.80 for 80%, 0.85 for 85%, 0.90 for 90%, 0.95 for 95%, 0.97 for 97% or 1.00 for 100%, and \bullet is the symbol for the multiplication operator, and wherein any non-integer product of x_a and y is rounded down to the nearest integer prior to subtracting it from x_a .

By way of example, a polypeptide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:2, that is it may be 100% identical, or it may include up to a certain integer number of amino acid alterations as compared to the reference sequence such that the percent identity is less than 100% identity. Such alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence. The number of amino acid alterations for a given % identity is determined by multiplying the total number of amino acids in SEQ ID NO:2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of amino acids in SEQ ID NO:2, or:

$$n_a \leq x_a - (x_a \bullet y),$$

wherein n_a is the number of amino acid alterations, x_a is the total number of amino acids in SEQ ID NO:2, y is, for instance 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc., and \bullet is the symbol for the multiplication operator, and wherein any non-integer product of x_a and y is rounded down to the nearest integer prior to subtracting it from x_a .

The terms “comprising”, “comprise” and “comprises” herein is intended by the inventors to be optionally substitutable with the terms “consisting of”, “consist of”, and “consists of”, respectively, in every instance.

“Immunogenic composition” in the context of a polynucleotide means that when the polynucleotide is introduced into a host and protein is expressed from that polynucleotide, the expressed protein is immunogenic.

"Individual(s)," when used herein with reference to an organism, means a multicellular eukaryote, including, but not limited to a metazoan, a mammal, an ovid, a bovid, a simian, a primate, and a human.

- 5 "Isolated" means altered "by the hand of man" from its natural state, *i.e.*, if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living organism is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. Moreover, a polynucleotide or
10 polypeptide that is introduced into an organism by transformation, genetic manipulation or by any other recombinant method is "isolated" even if it is still present in said organism, which organism may be living or non-living.

- "Polynucleotide(s)" generally refers to any polyribonucleotide or polydeoxyribonucleotide,
15 which may be unmodified RNA or DNA or modified RNA or DNA including single and double-stranded regions.

"Toxin" preferably includes a toxoid form of the toxin.

- 20 "Variant" refers to a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes
25 may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical.
30 A variant and reference polypeptide may differ in amino acid sequence by one or more

substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring
5 variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.

"Disease(s)" means any disease caused by or related to infection by a bacteria, including, for example, otitis media in infants and children, pneumonia in elderlies, sinusitis,
10 nosocomial infections and invasive diseases, chronic otitis media with hearing loss, fluid accumulation in the middle ear, auditive nerve damage, delayed speech learning, infection of the upper respiratory tract and inflammation of the middle ear.

EXAMPLES:

The examples below are carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described in detail. The examples are illustrative, but do not limit the invention.

Example 1: Cloning of the BASB232 genes *B. pertussis* strain Tohama I

Genomic DNA is extracted from the *B. pertussis* strain Tohama I from 10^{10} bacterial cells using the QIAGEN genomic DNA extraction kit (Qiagen GmbH). This material (1 µg) is then submitted to Polymerase Chain Reaction DNA amplification using two specific primers. A DNA fragment is obtained, digested by the suitable restriction endonucleases and inserted into the compatible sites of the pET cloning/expression vector (Novagen) using standard molecular biology techniques (Molecular Cloning, a Laboratory Manual, Second Edition, Eds: Sambrook, Fritsch & Maniatis, Cold Spring Harbor press 1989). Recombinant pET-BASB232 is then submitted to DNA sequencing using the Big Dyes kit (Applied biosystems) and analyzed on a ABI 373/A DNA sequencer in the conditions described by the supplier.

Example 2: Expression and purification of recombinant BASB232 proteins in *Escherichia coli*.

The construction of the pET-BASB232 cloning/expression vector is described in Example 1. This vector harbours the BASB232 gene isolated from *Bordetella pertussis* strain Tohama I in fusion with a stretch of 6 Histidine residues, placed under the control of the strong bacteriophage T7 gene 10 promoter. For expression study, this vector is introduced into the *Escherichia coli* strain Novablue (DE3) (Novagen), in which, the gene for the T7 polymerase is placed under the control of the isopropyl-beta-D thiogalactoside (IPTG)-regulatable *lac* promoter. Liquid cultures (100 ml) of the Novablue (DE3) [pET-BASB232] *E. coli* recombinant strain are grown at 37°C under agitation until the optical

density at 600nm (OD600) reached 0.6. At that time-point, IPTG is added at a final concentration of 1mM and the culture is grown for 4 additional hours. The culture is then centrifuged at 10,000 rpm and the pellet is frozen at -20°C for at least 10 hours. After thawing, the pellet is resuspended during 30 min at 25°C in buffer A (6M guanidine hydrochloride, 0.1M NaH₂PO₄, 0.01M Tris, pH 8.0), passed three-times through a needle and clarified by centrifugation (20000rpm, 15 min). The sample is then loaded at a flow-rate of 1ml/min on a Ni²⁺ -loaded Hitrap column (Pharmacia Biotech). After passage of the flowthrough, the column is washed successively with 40ml of buffer B (8M Urea, 0.1MNaH₂PO₄, 0.01M Tris, pH 8.0), 40ml of buffer C (8M Urea, 0.1MNaH₂PO₄, 0.01M Tris, pH 6.3). The recombinant protein BASB232/His6 is then eluted from the column with 30ml of buffer D (8M Urea, 0.1MNaH₂PO₄, 0.01M Tris, pH 6.3) containing 500mM of imidazole and 3ml-size fractions are collected. Highly enriched BASB232/His6 protein can be eluted from the column. This polypeptide is detected by a mouse monoclonal antibody raised against the 5-histidine motif. Moreover, the denatured, recombinant BASB232-His6 protein is solubilized in a solution devoid of urea. For this purpose, denatured BASB232-His6 contained in 8M urea is extensively dialyzed (2 hours) against buffer R (NaCl 150mM, 10mM NaH₂PO₄, Arginine 0.5M pH6.8) containing successively 6M, 4M, 2M and no urea. Alternatively, this polypeptide is purified under non-denaturing conditions using protocoles described in the Quiexpresssionist booklet (Qiagen GmbH).

Example 3: Production of Antisera to Recombinants BASB232

Polyvalent antisera directed against the BASB232 protein are generated by vaccinating rabbits with the purified recombinant BASB232 protein. Polyvalent antisera directed against the BASB232 protein are also generated by vaccinating mice with the purified recombinant BASB232 protein. Animals are bled prior to the first immunization ("pre-bleed") and after the last immunization.

Anti-BASB232 protein titers are measured by an ELISA using purified recombinant BASB232 protein as the coating antigen. The titer is defined as mid-point titers calculated by 4-parameter logistic model using the XL Fit software. The antisera are also used as the first antibody to identify the protein in a western blot as described in example 5 below.

Example 4: Immunological characterization: Surface exposure of BASB232

Anti-BASB232 proteins titres are determined by an ELISA using formalin-killed whole cells of *Bordetella pertussis* (*B.pertussis*). The titer is defined as mid-point titers calculated by 4-parameter logistic model using the XL Fit software.

Example 5. Immunological Characterisation: Western Blot Analysis

Several strains of *B.pertussis*, as well as clinical isolates, are grown on Bordet Gengou agar plates for 24 hours at 36°C and 5% CO₂. Several colonies are used to inoculate Tryptic Soy Agar (TSA) broth supplemented by NAD and hemin, each at 10 µg/ml. Cultures are grown until the absorbance at 620nm is approximately 0.4 and cells are collected by centrifugation. Cells are then concentrated and solubilized in PAGE sample buffer. The solubilized cells are then resolved on 4-20% polyacrylamide gels and the separated proteins are electrophoretically transferred to PVDF membranes. The PVDF membranes are then pretreated with saturation buffer. All subsequent incubations are carried out using this pretreatment buffer.

PVDF membranes are incubated with preimmune serum and rabbit or mouse immune serum. PVDF membranes are then washed.

PVDF membranes are incubated with biotin-labeled sheep anti-rabbit or mouse Ig. PVDF membranes are then washed 3 times with wash buffer, and incubated with streptavidin-peroxydase. PVDF membranes are then washed 3 times with wash buffer and developed with 4-chloro-1-naphtol.

Example 6: Presence of Antibody to BASB232 in Human Convalescent Sera

Western blot analysis of purified recombinant BASB232 is performed as described in Example 5 above, except that a pool of human sera from children infected by *B.pertussis* is used as the first antibody preparation.

5

Example 7: Efficacy of BASB232 vaccine: enhancement of lung clearance of *B.pertussis* in mice.

This mouse model is based on the analysis of the lung invasion by *B.pertussis* following a standard intranasal challenge to vaccinated mice.

- 10 Groups of mice are immunized with BASB232 vaccine. After the booster, the mice are challenged by instillation of bacterial suspension into the nostril under anaesthesia. Mice are killed between 30 minutes as well as 2, 5 and 8 days after challenge and the lungs are removed aseptically and homogenized individually. The log₁₀ weighted mean number of CFU/lung is determined by counting the colonies grown on agar plates after
- 15 plating of dilutions of the homogenate. The arithmetic mean of the log₁₀ weighted mean number of CFU/lung and the standard deviations are calculated for each group. Results are analysed statistically.

In this experiment groups of mice are immunized either with BASB232 or with a killed whole cells (kwc) preparation of *B.pertussis* or sham immunized.

20

Example 8: Useful Epitopes

- The B-cell epitopes of a protein are mainly localized at its surface. To predict B-cell epitopes of BASB232 polypeptides two methods were combined: 2D-structure prediction and antigenic index prediction. The 2D-structure prediction was made using
- 25 the PSIPRED program (from David Jones, Brunel Bioinformatics Group, Dept. Biological Sciences, Brunel University, Uxbridge UB8 3PH, UK). The antigenic index was calculated on the basis of the method described by Jameson and Wolf (CABIOS 4:181-186 [1988]). The parameters used in this program are the antigenic index and the minimal length for an antigenic peptide. An antigenic index of 0.9 for a minimum of 5

consecutive amino acids was used as threshold in the program. Peptides comprising good, potential B-cell epitopes are listed in table 5. These can be useful (preferably conjugated or recombinantly joined to a larger protein) in a vaccine composition for the prevention of Bordetella infections, as could similar peptides comprising conservative mutations (preferably 70, 80, 95, 99 or 100% identical to the sequences of table 5) or truncates comprising 5 or more (e.g. 6, 7, 8, 9, 10, 11 or 12) amino acids therefrom or extensions comprising e. g. 1, 2, 3, 5, 10 further amino acids at either or both ends from the native context of BASB232 polypeptides which preserve an effective epitope which can elicit an immune response in a host against the BASB232 polypeptides.

Table 5: Potential B-cell epitopes from SEQ ID NO:34

Position	Sequence
56	QDAGQEGEF
84	DDDPDELGE
106	EHKNPMS
236	PGFPPPPPLP
265	GQDGK
339	DGANT
381	TLRQTRI
404	PQSGSG
538	DGNKPL
579	ADSRVQD
599	APEASY
628	QNDQL
636	GRADGQ
653	ADSRGA
692	AEDPKT
753	TFSERQQISNRH
766	RAYDQT
785	ASGGRW
800	YADRTYPGDGGG
839	GRYDQQY
858	DYRTSG
869	EGGRRF
893	TSGKRYRASN
944	QEFKSTGDVRTNG
962	AGRHGR
1004	AGYRYSF

The T-helper cell epitopes are peptides bound to HLA class II molecules and recognized by T-helper cells. The prediction of useful T-helper cell epitopes of BASB232

polypeptides was based on the TEPTOPE method describe by Sturniolo et al. (Nature Biotech. 17: 555-561 [1999]). Peptides comprising good, potential T-cell epitopes are

listed in table 6. These can be useful (preferably conjugated to peptides, polypeptides or polysaccharides) for vaccine purposes, as could similar peptides comprising conservative mutations (preferably 70, 80, 95, 99 or 100% identical to the sequences below) or truncates comprising 5 or more (e.g. 6, 7, 8, 9, 10, 11, 12, 14, 16, 18 or 20)

- 5 amino acids therefrom or extensions comprising e. g. 1, 2, 3, 5, 10 further amino acids at either or both ends from the native context of BASB232 polypeptides which preserve an effective T-helper epitope from BASB232 polypeptides.

Table 6: Potential T-helper cell epitopes from SEQ ID NO:34

Position	Sequence
20	WRLHALAAALA
34	MARLAPAAA
105	VEHKNPMSK
118	VRVSGAGRA
144	VVRGGTLELDGVTVA
164	MEPMTVSDA
192	LVRAAQGGQ
208	LQSILGPALADGGSIS
270	VTLREVALRAHGPOAT
287	VYAYMPGSEI
298	LQGGTVSVQ
329	VRLDGTTVS
347	LVRGDAARAEEVNTVLRTAKSLAA
380	VTLRQTRIE
420	ITTTGNRAA
444	VRAEGSGSS
461	LVVSAGSLAS
483	LKLMPGALASS
497	VRLTDGATA
513	LQQHSTIPV
535	IVADGNKPL
563	VLQSATLGK
576	VVNADSRVQD
586	MSMRGGRVEFQAPAPE
609	LQTLDGNGVFVLNTNVAA
644	VLVRNAGGEA
660	LGLVHTQGG
673	FRLANVGKAVD
687	WRYSLAEDP
699	VWSLQRAQALS
725	IALAESNAL
740	LRLRADAGGPWARTFSERQ
824	YVGDGGYYLDTVLRLGRYDQ
845	YNIAGTDGG
859	YRTSGAAWS
888	VMLWRTSGKRYRASNGLRV
915	LGRLGLRFGRRIALAG
934	VQPYARLGW
952	VRTNGIGHA
996	INIPWSFHA

The same analyses (B-cell epitopes prediction and T-helper cell epitopes prediction) could be done for each BASB232 polypeptide sequence comprised in SEQ Group 2.

Example 9: Expression of BrkA

BrkA is a 1010 amino acid protein. A pET30b expression vector containing amino acids 60-702 of BrkA, fused at the C and N-terminal to a 6x-His tag was used to express BrkA in *E. coli*. The bacteria were grown to an OD600 of approximately 0.6 and induced with 1mM IPTG for 2 hours. Recombinant BrkA was purified under denaturing conditions using the protocol in the Xpress System Protein Purification manual (Invitrogen, Carlsbad, CA). The bacteria were lysed in 6M guanidine hydrochloride and the lysate was applied to to Ni²⁺ -nitrilotriacetic acid agarose (Qiagen, Mississauga, Ont.). After successive washes in 8M urea of decreasing pH, purified BrkA was eluted at pH 4 and the fractions were pooled. The urea was removed by slow dialysis at 4 °C against 10mM Tris, pH 8.0 in the presence of 0.1% Triton X-100.

The purified protein was refolded so that it has a beta-structure resembling PRN (monitored by CD) and was functional in a serum assay. The protein was filter sterilized and is diluted in 10mM Tris buffer, pH 8 at a concentration of approximately 0.4mg/ml.

Example 10: Formulation of vaccines

Six vaccines were formulated:

1. DTPa018A2 – contains 6.25Lf DT, 2.5Lf TT, 6.25ug PT and 6.25ug FHA per mouse dose. All antigens were adsorbed separately onto Al(OH)₃ before combining.
2. DTPa14885B9 - contains 6.25Lf DT, 2.5Lf TT, 6.25ug PT, 6.25ug FHA and 2ug pertactin per mouse dose. . All antigens were adsorbed separately onto Al(OH)₃ before combining.

3. DTPw13126A9 – contains 6.25Lf DT, 2.5Lf TT and 1 I.U.whole cell *B. pertussis* per mouse dose. All antigens were adsorbed separately onto Al(OH)₃ before combining.
4. DTBrkA - contains 6.25Lf DT, 2.5Lf TT and 2.5ug BrkA per mouse dose. All
5 antigens were adsorbed separately onto Al(OH)₃ before combining.
5. DTPa-2 BrkA - contains 6.25Lf DT, 2.5Lf TT, 6.25ug PT, 6.25ug FHA and 2.5ug BrkA per mouse dose. . All antigens were adsorbed separately onto Al(OH)₃ before combining.
6. DTPa-3 BrkA - contains 6.25Lf DT, 2.5Lf TT, 6.25ug PT, 6.25ug FHA, 2ug
10 pertactin and 2.5ug BrkA per mouse dose. All antigens were adsorbed separately onto Al(OH)₃ before combining.

Example 11: Protection against lung invasion by *B. pertussis* in an animal model

- 15 Groups of 20 BALB/c mice (females, 5 weeks old) were immunized subcutaneously with ¼ of a human dose (125 µl of vaccine) and were boosted 3 weeks later. One week after the booster, a sample of blood was collected from each mouse for antibody determination. The mice were then challenged by instillation of 50 µl of bacterial suspension (+/- 5 10⁶ CFU/50 µl) into the left nostril under ether anesthesia. Five mice
20 in each group were killed at 4 different times (2 hours, 2, 5 and 8 days) after challenge and the lungs were removed aseptically and homogenized individually. The log₁₀ weighted mean number of CFU/lung was determined by counting the colonies grown on Bordet-Gengou agar plates after plating of 100 µl of 4 serial dilutions of the homogenate. The arithmetic mean of the log₁₀ weighted mean number of CFU/lung
25 and the standard deviation were calculated for each group and each time point.
- Three experiments were performed. The same vaccine groups were included in each experiment but animals were either challenged with *B. pertussis* strain Tohama, *B. pertussis* strain 18323 or *B. parapertussis*.

The day before challenge, blood was collected from each mouse. The anti-PT, anti-FHA, anti-PRN and anti-BrkA antibody levels were determined by ELISA. The geometric mean titre for each group of mice was calculated.

5 **Example 12: *B. pertussis* strain Tohama challenge**

In one experiment immunized mice were challenged with *B. pertussis* strain Tohama. The number of CFU/lung at each timepoint after challenge and for each group is summarized in the figure 1. The experiment had low variability with the data having a mean square error of is 0.450.

- 10 Statistical analysis of the data using ANOVA1 was used to assess the data. No significant difference was seen between the protection against *B. pertussis* offered by DTBrkA compared to control, indicating that immunisation with BrkA alone is insufficient to elicit protection. In contrast, the addition of BrkA to a DTPa-2 vaccine produced a statistically significant increase in protection showing that, in combination
- 15 with PT and FHA, BrkA can produce additional protection. The level of protection conferred by DTPa-2 BrkA was statistically slightly less than that conferred by DTPa-3 which conferred protection statistically equivalent to DTPw. The DTPa-3 BrkA vaccine provided excellent protection from challenge after 2 and 5 days but less protection after day 8.

20

Example 13: *B. pertussis* strain 18323 challenge

In this experiment, immunized mice were challenged with *B. pertussis* strain 18323. The number of CFU/lung at each timepoint after challenge and for each group is summarized in figure 2. The experiment showed low variability with the mean square error of the experiment being 0.402. Statistical analysis using ANOVA1 showed that

25 again DTBrkA did not provide a significant protection over the control. However, DTPa-2 BrkA provided better protection than DTPa, showing that BrkA, in

combination with other *B. pertussis* antigens gives additional protection. The protection achieved by vaccination with DTPa-2 BrkA, against challenge with *B. pertussis* strain 18323, was statistically equivalent to that provided by DTPa-3, DTPa-3 BrkA and DTPw.

5

Example 14: Comparison of protection against *B. pertussis* and *B. parapertussis* in mice vaccinated with DTPw or DTPa

Groups of 25 or 30 BALB/c mice (females, 5 weeks old) were immunized subcutaneously with ¼ of a human dose of DT, DTPa or DTPw from different sources (125 µl of vaccine) and were boosted 3 weeks later. The sources of DTPw were Triple antigen (CSL), Tri-immune (Lederle), Pentacoq (MSD), Combivax (Behring), Infanrix (SB), DKTP (RVM), DTPw (Connaght) and Trivax (Wellcome). The sources of DTPa were Infanrix (SB), Triacel (PMCS), DI-TE-KIK (Amvax), Acell-immune (Lederle), Tropedia (Biken), Tricelluvax (Biocine/Chiron), Pentavac (PM-MSD) and DTPa-2 (SB). One or two weeks after the booster, a sample of blood was collected from each mouse for antibody determination. The mice were then challenged by instillation of 50 µl of bacterial suspension ($\pm 5 \times 10^6$ CFU/50 µl) into the left nostril under anaesthesia. Five mice in each group were killed at 5 or 6 different times (ranging from 2 hours to 14 days) after challenge and the lungs were removed aseptically and homogenized individually. The log₁₀ weighted mean number of CFU/lung was determined by counting the colonies grown on Bordet-Gengou agar plates after plating of 100 µl of 4 serial dilutions of the homogenate. The arithmetic mean of the log₁₀ weighted mean number of CFU/lung and the standard deviation were calculated for each group and each time point.

The day before challenge, blood was collected from each mouse for determination of the anti-PT, anti-FHA and anti-PRN antibody levels by ELISA. The geometric titre for each group of mice was calculated.

Results were analysed statistically by applying 1-way and 2-way ANOVA after assuming equality of variance (checked by Brown and Forsythe's test) and normality (checked using the Shapiro-Wilk test). Differences between groups were analysed using the Dunnet test, Tukey's studentised range test (HSD) and Student-Newman-Keuls test.

Results are shown in figure 3, and show that the DTPw vaccines induced good lung clearance of *B. pertussis* and *B. parapertussis*. However, DTPa vaccines induced strong lung clearance of *B. pertussis* but only limited clearance of *B. parapertussis*.

This experiment demonstrates that antigens present in the DTPw vaccine and absent from the DTPa vaccine are protective against parapertussis. Moreover, the antigens present in the DTPa vaccine, specifically FHA and pertactin, may display antigenic variability when compared to the corresponding parapertussis antigens.

Deposited materials

A deposit of strain 3 (strain 3224A) has been deposited with the American Type Culture Collection (ATCC) on May 5 2000 and assigned deposit number PTA-1816.

5

The *B. pertussis* strain deposit is referred to herein as "the deposited strain" or as "the DNA of the deposited strain."

The deposited strain contains a full length BASB232 gene.

10

The sequence of the polynucleotides contained in the deposited strain, as well as the amino acid sequence of any polypeptide encoded thereby, are controlling in the event of any conflict with any description of sequences herein.

15 The deposit of the deposited strain has been made under the terms of the Budapest Treaty on the International Recognition of the Deposit of Micro-organisms for Purposes of Patent Procedure. The deposited strain will be irrevocably and without restriction or condition released to the public upon the issuance of a patent. The deposited strain is provided merely as convenience to those of skill in the art and is not an admission that a deposit is required
20 for enablement, such as that required under 35 U.S.C. §112. A license may be required to make, use or sell the deposited strain, and compounds derived therefrom, and no such license is hereby granted.

SEQUENCE INFORMATION

BASB232 Polynucleotide and Polypeptide Sequences

- 5 SEQ group 1 contains SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 35, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95 and 97.

- 10 SEQ group 2 contains SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96 and 98.

SEQ ID NO:1 polynucleotide sequence of Orf1

Atgtccaccccccgattcgcgctgcattacgccagcgcgctcagtcctgctggccgcgcatcc
 15 Ggcctggccatggcgcgagacggccacccagatccacgatccgtcgcgaggtgcagcagatg
 Gcgacggtgcaggtgctgggacggccgaagaggaaatcaaggagtcgctgggctctcg
 gtcatcacccgagaggagatcgcccgcggcccccacaaatgacctgtccgacctgatc
 cgccgcgaacccggcgtaacctgacgggcaacagcgccagcgcgcgcggggcaacagc
 cgccaggtcgacatccgcgccatgggccccgagaacaccctcatcctgatcgacggcaag
 20 cccgtcacctcgcgcaatgcggtgcgctatggctggaacggcgaccgggacacgcgcggg
 gacaccaactgggtgcccgcgaggaagtgcgagcgcatcgaagtgatccgcgggcccgcc
 gccgcccgtacgggtccggggccatgggcgcggtggtcaacatcatcaccaagcgcccc
 gccgatcgcgccacgggtccatcacctactacacgaaccagccggaagacagcccgag
 ggcaacaccaaccgctcaatgcgcgcacatcagcgcgccgatcagcgacacgctgagcatg
 25 cggctgtacggcaactacaacaagaccaatccggatgcccgcgacatcaacgccggccac
 gcgaacaccagcgacaacggcaacccctcgaccgcccagcgagggcgctcatcaaccag
 gacctgagcgcgctgttctcgtggaaagccgacagccacaacaccgtggacctggacatg
 ggcttcagccggcagggcaacctgttcgcccgcgacaccatgaacaacgccaacagcgac
 30 ttctcggacagcctgtacggcaaggaaaccaatgcgatgtaccgcgagaactatgcgctg
 acgcaccgcggcgctctacgactggggcacctcgcgcgccagcgctcggtatgactacacg
 cgcaacgcgcgcgcagcgcggaaggcctggccggcgggcccgagggcgcgccaccgcgggc
 ggctacgacacccgcgcgcctgaagaactggcgcgcccgcgccgagggcagcgctgccgttc
 catctcgggttcgagcaggtcgccacggtcggcggtggaatggctgcgcgaatcgctggaa
 35 gaccccgccggcacgcgcgcagaccagctatcgcgctgttcgcccaggacaacatcgag
 atcgacgagcgcacccatgctcacgcccggcgctgcgcctggaccacaacagcgaattcggc
 agcaactggagtcaccagcctgaacgcctcgtacgcctcaccgacgcgctcaagctcaag
 ggtggcatcgcgcgcgccataaaggcgcccaacctctaccaatccaaccccaactacctg
 40 ctgtacagccgcggcaatggctgcctggcctcgagaccaacaccaacggctgctatctg
 gtgcggaacgaggacctctcgccggaaccagcgctcaacaaggaaatcggttcgagtac
 gacccgggacgctggcgacaccagcatggcctatttcgcaacgactaccgcaacaagatc
 gtcgcccggcacccgacgtccagtagccgctggccaatggcgccccgggtgctgcaatggacc
 aacagcggaaggccgtggctcgaagggtggaaggcaacctgttcattccgctggccagc
 aatctcgactggaacaccaacttcacctacatgatccagtcgaaggaaaaggctaccggc
 45 gaaccttgagcgtgattcccgaatacaccatcaacagcacgctggactggttctacacg
 ccgcagctgtcgttcaggccaatctcacctattacggcaagcaggaaggcccgctccacc
 aatgtacgcacccggcgtaactgaacggcgacggccgacccatcagtcctgatgcc
 ctggcgggcctgagcatgggctacgaagtcaaccggaacctgaagtcccgctcggcgtg
 agcaacctgttcgacaagcagctgtaccggaaggcaatgccagcagcgcgggcgcgcc
 acctacaacgaaccggggcgcgccatattacgccacggcgacggtgtcgttctga

SEQ ID NO:2 polypeptide sequence of Orf1

5 MSTPRFALHYASASVLLAASGLAMAQTATQIHDPQSQQMATVQVLGTAEKEEIKESLGVS
 VITAEIARRPPTNDLSDLRREPQVNLTGNSASGARGNSRQVDIRGMGPENTLILIDGK
 PVTSRNAVRYGWNDRDTRGDTNWVPAEEVERIEVIRGPAAARYGSGAMGGVNNIITKRP
 10 ADRTGSGITYYTNPQEDSREGNTNRVNARISAPISDTLSMRLYGNYNKTNPDARDINAGH
 ANTSDNGNPSTAGREGVINQDLSALFSWKADSHNTVDLDMGFSRQGNLFAGDTMNNANS
 FSDSLYGKETNAMYRENYALTHRGVYDWGTSRASVGVDYTRNARQREGLAGGPEGAPTAG
 GYDARLKNWRAAAEASVPFHLGFEQVATVGVEWLRESLEDPAQTRQTYTGGAIIGGTAPA
 15 DRDPKSRQTSYALFAEDNIEIDERTMLTPGVRLDHNSEFGSNWSPSLNASYAVTDALKLK
 GGIARAYKAPNLYQSNPNYLLYSRGNGCLASQTNTNGCYLVGNEDLSPETSVNKEIGFEY
 DPGTWRTSMAYFRNDYRNKIIVAGTDVQYRLANGARVLQWTNSGKAVVEGLEGNLFIPLAS
 NLDWNTNFTYMIQSKEKATGEPLSVIPEYTNSTLDWFYTPQLSFQANLTYYGKQEGPST
 NVRTGVELNGDGRQTIISPYALAGLSMGYEVNRNLKFRVGVSNLFDKQLYREGNASSAGAA
 TYNEPGRAYATATVSF

SEQ ID NO:3 polynucleotide sequence of Orf2

20 Gtgttttctcgagtcagaagcatccgtcctggcgccctgtcgccatgcgtacttgcggcc
 Gccttgtgcgccgtcgcggtcggtagcgccgacacccgcccgcgcgcagggcgcccgccg
 Agcgcccagcattatgaaatcgcgccgacccgctggccgacgcactgacccgcttcgcg
 cgccgtgcccgcgtggtgctgtcgttcgacccggccctgggtgcagggcgccgacgagc
 ggccgtgcagggcggtgtacggcggtgcgcgacgggttcgcgccgctgctggccggctcg
 ctgcagggcgccgcccggcgccgaacaactggctcgctggcgccgctgcccgcgcggcggc
 gatgcgcagacgctggcgccggtgacgggtgctgggcctggagggcgcgctggcgcccg
 25 gtcggctatgtcgccagtgccagcctgagcgccaccaagaccgatacgcgcgtgatcgaa
 acgcccgaatcgatttcgggtggtgactcgcgaccagataaccgagcagggcgcccgacg
 ctgaaccaggtgctgcgctataccgcccgtggcgaccgagacgcgcggcgccgacgagc
 acgcccgtggacaggttcagcgtgcgcggtttctccgcccacacatctggacggcatg
 cgcggttccggcgccgggacgccttgcccaggtcgacgcctaccggctcgaaacgggtc
 30 gatgtgctcaagggcgccgcttcggtgctgtatggccagggcgcccgggcgccggtc
 aaccaggtcagcaagcgtcccctggacgagcctttgcgcgagatcgaaagtcagggcgcc
 aatttcgatttcggcgccgggtcaacatggatttttcggcccgggtggacgaggaacggcg
 ttccgtgacgggtaaccggcgccgctatatgtccgatggccaggtggatcacaccagg
 gagcgccgctacttcgtctcgccgctcggttcacgtggcgcccgccagcgccgataccacg
 35 accgtgctgaccaacttcagcgccgaccccgacatgggttcgtacggatcgatctcgcc
 atgcgcacgctgctgtcgccgcccagcgccagggcggtgggcccgaaccactacgacggc
 gacggcgatttcgaaaagagcgacccgcccagcagctattcgctgggctatcaactggagc
 cgcttcaacgataccttcaaggcctcgcaaacctgcgtttccagcatgccgagggcgctc
 tatcgagcatctacggcgccagcaacaacaattacggctatctcgacaaggactaccgc
 40 tactcgacgcgccgctggccatcagcgacgtggacgtggatgcgttcacgatcgacaac
 aacctgcagggcgcccttgataccggggcgctggcgccatcggtgctgggtggggttcgac
 taccagcgctgcagaccgacaccttgcgggctatggcagcgccgcccgcgtcgacgtg
 ttcgatccggactatcacatgggtatcgagcgcccgccgctttacgtccgatcagacccag
 tacaactaccagaccggcctctacctgcaggaccagatcaggctggatcgccgtgctgtg
 45 ctgctggcgccgctacgactggtcgcgcccccacacgggcacccgacaacctggccaac
 ggctcgacagcagctcgccgctcgccgcccagggcgttcacccggccgggtcgggcgatc
 tacaacttcgacaacggcggtggcgccgtacgccagctactcgagtcgttcgagccgag
 accggcacgggctggaaacaacacgcccgttcaagccgaccgaaggcaagcagatgaggtc
 ggcgtgaaataaccagccgcccgggctcgccacacgctgctcacgctggcgcccttcgacatc
 50 cgccgcaagaacctgcccaccacggccggtatccacccatattgtcgccggttttcgccc
 tgctcgatacagggccggcggaagtgcgcacccgcccgcacgaactggaggccaaagaccga
 ccgctgcgcggcctgagcctgatcgccgcctattcgtacctcgacaacgaatacgaagaag
 gcctatccgaacacgaccgggttgacacacagggaagaagccggtggccgctggccg
 caccagcgctcgccctgggcccgtatcaactgcaggaggcccgctggccggcctggg
 55 atggcgccgggggtgcgctacatcgccaggttcgtacgccaacgaaaccaacacgctcaag
 gtgccatcggtgacgctgggtggacatgatgctcgactacgacacggggccggccagcccc
 gcgctcaaggccatgcaggtggcggtgaacgtctccaacctgttcgacaaggaatacatc
 ggctcgtgcctgtccgattcgtgggtgctgggtatgggtaccagcggttcgatcaaggccagc

ttgcgctatcgctggtga

SEQ ID NO:4 polypeptide sequence of Orf2

5 VFSRSQKHPSWRLSPCVLAAALCAVAVGSADTARAQAPAASAQHYEIAAGPLADALTRFA
 RRAGVVLSPALVQGRSTAGLQGVYGVDRDGAALLAGSGLQARAGGNNWSLAALPRGG
 DAQTLAPVTVLGLLEGALPTVGYVASASLSGKTKDTPLIETPQISVVTTRDQITEQGAQT
 LNQVLRITAGVATETRGATATRLDQFSVRGFSAAATYLDGMRVFGGRDALPOVDAYRLERV
 DVLKGPASVLYGQGGPGGVNVQVSKRPLDEPLREIEVQAGNDFRRVNMDFSGPVDEDRR
 10 FLYRVTGAAYMSDGQVDHTRRERYFVSPSFTWRPSADTTLTVLTNFORDPDMGSYGSISA
 MRTLLSAPDGRRLGPNHYDGDADFEKSDRRSYSLGYQLEHRFNDTFKASQNLRFQHAEGV
 YRSIYGASNNNYGYLDKDYRSQRGLAISDVVDVDAFTIDNNLQARFDTGALAHTVLVGF
 YQRVQTDTLGSGYSAPPLDVFDPDYHMGIERPPFTSDQTQYNYQTGLYLQDQIRLDRLSL
 LLGGRYDWSRTHGTGTDNLANGSHSSSALAAEAFTGRVGAIYNFDNGVAPYASYSESFEPQ
 TGTGWNNTPFKPTGEGKQYEVGVKYQPPGSATLLTLAAFDIRRNLPPTDPTDPTMCGVSR
 15 CSIQAGEVTRTGIELEAKTEPLRGLSLIAAYSILDNEYEKAYPNTTGLDLKGKKPVAVPA
 HQASAWARYQLQEGPLAGLGMGAGVRYIGSSYANETNTLKVPSTLVDMMLDYDLGRASP
 ALKGMQVALNVSNFLDKEYIGSCLSDSWCWYGYQRSIKASLRWR

SEQ ID NO:5 polynucleotide sequence of Orf3

20 Atgcgagaccagaccgagcgcggaacccgcgcctgcactcatcgctcgcccccgccgc
 Ctgatccctgccctgctggcgcgcttgccctgctgggcactggcgctacaggccgcgccc
 Atcgacgtcgatatcccgccccagaacctggcccaagccctgcaccagcttggccggcga
 gccaacctgcaggtcctgtacagccaggacctggtcgatggccagcgagcccccgccgtg
 cagggcccgcatggaacctgcccgaagcgctggaacgcctgctgaaaggccgaacatccgc
 25 tattcgatccagcacaacaccgtcacgctcacgcccattgccgtgactgcgacgctgccc
 gcaatcagcgctggcgccgcccctgctgactccgacacctacgtggccacaggcacgaca
 gccggccaccaagacggacacgcgcgctgatcgaaataccgcaatccatttcgggtggtgacc
 gccgcgcgagatccgcgagcagaatccgcgacgctggcgacgcccgtgctgctacacgccc
 ggtatcggtggtgcaggaaggattcaaccgcaccgacgatcccttcacatcccgcggttc
 30 gacgtccgcaccaatcccgcgctcatgttccgcgacgggctgaaaaatccccctgccccat
 tacagcgcgatgtccgaaccctacgcgctcgaacgcacgaggtcgtaaggcccccgt
 tcggtgctgtgagcagggcctcgccaggcgcatcgtaacgtggtatccaagcgccg
 acagacagccccgctgcgcgagctgcagctgagcgccggtcgacacagcaacaggcagctc
 gccggcgacttcggcgagcgcacgacgagggacggctgacctaccgcctgaccggc
 35 ctggcgcgcaatgccgacacgatgatcgacctgtcccgacgacgctactatctcgcc
 cccgctgacctggcgcatcagcccgacacctcgctgacgctgctggcgagctacatg
 aagaacaagaccatcaacaacgcccgtaccgctcgaaggcacgggtcaagtaacaaccc
 aacggccgcatcccgctcaccgcttcacggcgagccggactggagcaagtgggaccag
 gaggtcgccaacgtggggtaccagtttgcgcaccgcttcaacgacacctggcaattcaag
 40 cagaacctgggctacgcccagtcgcgcaaccgctcaaccacgcctactggtggacctgg
 gtgcccgaagcgacttctccacggcgagcgccgctaccgcccgcgacgatgacgcc
 cagggcgtgagcatcgacaaccagttcgaggccacgtggcaatccggccgcttcaggcac
 aacacgctgttcggcctcgattacaccgaaacctcggtcaccgcaaacagtagcggc
 tacaacaacctcgccccgatcgacttcttcgatccggcgtagcgctcggacgtactgctg
 45 ccggcgaagccggacacctacaccaacgagaagcgagccagctcgccctgacttgacg
 gaccagataaagttcgacgacaaactgggtggtggtgctcagcgccgctacgacaatgcc
 gacggctcgacgctgaacaagctgagcgccgctcaatacccgacccggcgacaacgcgttc
 acgtggcgacccggcctgctctaccttgcggacaacggcttggcgccctataccagctat
 tcgacttcgttcagcccgagccggcaccacctcgcccgacggcgccacacgccccttc
 50 gacccgaccaagggcaagcagtggaagccggcggtgaagtaccagcccaatgggttcgaac
 tcgttcacacccgcatccgtcttcgagctgacgcgcaccaacgtccccacgaccgacccc
 gccaaccccgctctacagcgtgcaggaaggcgaggtgcgctcgccggcgccctcgaaattgctg
 gccaccgccaacctggcctcgggctggaacctgatcgcgccctacacgtataccgacgcg
 gaaatcaccaagagcaactccaacacgctaggcaacacgcccagggccgtgcccgcgaac
 55 atggcgctgctatggtccgactacacgcctcccgctccgggtgcgctggcggggctgaatc
 ggcgcggcgctgcgctacatgggctcgacctacaacaacaccaatgccgccaaggtcggc
 gactacacgctgttcgacgcccgcctcgctacgacttcggggcgcgagccccgctccctg

aaaggctggacggccgatctcaccgtgcgcaacctgttcgacaaggactacgtggcctcg
 tgcacctatgcctgcttctacggagaaggcaggaccgtgctgggcccgggtcacgtacaaa
 tggtag

5 **SEQ ID NO:6 polypeptide sequence of Orf3**

MRARPSAATRALHSSSRPRRLIPALLGALACLTGTGVQAAPIDVDIPPQNLAQALHQLGRQ
 ANLQVLYSQDLVDGQRSPAVQGRMEPAEALERLLKGRNIRYSIQHNTVTLTPMPLTATLP
 AISVVGALPDSPTYVATGTTAGTKTDTPLIEIPQISVVTAAQIREQNPTLGDVRYTP
 GIVVQEGFNRTDDPFIIRGFDVRTNPGVMFRDGLKIPLPHYSAMSEPYALERIEVVKGPA
 10 SVLYGQASPGGIVNVVSKRPTDSPLRELQLSGGSHSNRLAGDFGGRIDDEGRITYRLTG
 LARNADTMIDHVPDDRYYLAPALTWRISPDTSLTLLASYMKNKTINNAGYPLEGTVKYNP
 NGRIPRHRFTGEPDWSKWDQEVANVGQFAHRFNNTWQFKQNLGYAQSRNRVNHAYWWTW
 VPGSDFSTAERGAYRRDDDAHGVSIDNQFEATWQSGRFRHNTLFGLDYTETSFTRKQYAG
 YNNLAPIDFFDPAYGSDVLLPAKPDITYNEKRSQGLYLQDQIKFDDKLVVVLSGRYDNA
 15 DGSTLNKLSGVNTRTGDNAFTWRTGLLYLADNGLAPYTSYSTSFQPOAGTTSPARGTTPF
 DPTKGKQWEAGVKYQPNGSNSFITASVFELTRTNVPTTDPANPVYSVQEGEVRSRGLELS
 ATANLASGWNLIAAYTYTDAEITKSNSNTLGNTPAVPRNMASLWSDYTVPSGALAGLNI
 GAGVRYMGSTYNNNTNAKVGDTVFDAALRYDFGARSPSLKGWTADLTVRNLFDKDYVAS
 CTYACFYGEGRTVLGRVITYKW

20

SEQ ID NO:7 polynucleotide sequence of Orf4

Atgacaggctttcatgcacgcaagccggtgggcccgtggccatggccggcccgcgcacggc
 Aggcgcttgcggtggccgcgcacccccctacggacaaccaccatgaagccattaccgctc
 Gcttatctcgccgcgctgctgccctggtacgcaggcgctcatccaggcgcaatccgcgcgc
 25 gccgcggcgacgatgcctcgatcaccctggaagccgtcagggtcgaggccagcgccgac
 gcctccgcggcgccgctggcgccggccttcgcggggcgccagggtcgccacggcgcggaag
 gtcggcatcctcggcacgcgcgacaacctggaaaccccgcttctccatcaccgcctacacc
 aacgaactgatccaggaccgcccaggccaaggggggtgggcgacgtcctgcagaacgacccc
 ggcggtgcgggtggcgcgccgggttcggcaacttcaggagtcgtatttcatccgcggcttc
 30 atcctcagctcggacgacatcgcggtacaacggcctctatggcctggtgccgcgccagtac
 atctcgacccagctgttcgagcgcgctcgagggtgctgcgcggtgcctcgcgcttctcacc
 ggcgcgcccgcgctccggcgccgggatcgccgggggtgatcaacctgggttcccaagcgcgcg
 cccaacgagccgctgacgcgcttttcggccgggtacggcgagcgacagcggtgctcgaggct
 tcggccgacatcgcccgccgcttcggcccgacgacagcgctcgggatccgcacacggcc
 35 gccagcgccggcgccgagacggccatcgacggcgagcgaccccgaccacgggtgttcgcg
 ctgggctggaactggcgcgccgagcgcgcgccgtgctggccgatatcggtaccaggac
 aaccgcctgaagcgggcgcccccattgtcacgctggccggcgacgcgcgccaaggtgcc
 ggcgcgcccgacgcccggctccaactatgccagccctggctcgactccaacgaacgagac
 gtgttcggcaccctgcgcggcggaatacgaactcaacggccgcataacgggctgggtcgcc
 40 tatggcatgcgccagagcaaggaggagaactcgctggccaacctcaataacgtcaacggc
 gcggggcagggcaagttctaccgcttcgacaacgcccgcgaggataccgtcaacaccggc
 gagatcgccctgcgcgccaaggcgcgacccggccgggtgggccaagaaactggctcgctcg
 gcgtcgatttccgacctcgagaagaagaacgcctatgtcatggacttcttcaaccagttc
 gacaccagcatctacgaccccgctcagctacgccaagccggcctcagcagcaccgcgctt
 45 cgccggaacgacatggacgatcccggccaagcaggcgctcatccggctggccagctatgcg
 ctggcgacacccatgtcgcttcttcgacgacaaggtgctgctgaccgcccggcatccggcac
 cagcgctctaccagcgcgactacagctacgacacgggcatcgccggcacccttacgag
 caaagccacaactcgcccgccggccgctgggtggcgctgacgccccaggtgtcgctg
 tacgccaactacatcgaggccctgtcgccggcgacacccgcgcccagaccgccaacggc
 50 ctgcccgggtggtaaccacggcggaatcgctggcgccctggcgctcttttccaccgacaag
 ggcgtcaagttcgagcacgacggcctggcgccggcctggcgctcttttccaccgacaag
 ccgcgcccgggttcgtggcgatgaccaggtcttccgcgcttcgggcaaggaccgcccaccgc
 ggggtcgaaactgacgacttacggcgagctcacgcgcagcggtgcgcgctgctggcgggctg
 acctggctggacgccaagcagctcagcaccggcaacgcccaccgacggcgaagcgcgctc
 55 atcgccgtgcccgccttccaggccaacctcgccgtggagtgaggacatccccggcgctgcag
 ggcctgaccgtggacggcgctggtctatacgggctcgctcctatgcggatgcggccaac
 accctcgaggtgcccgggctggacgcgctggacgcggcctgcgttacatgaccgatatac

ggcgcccatctggtgacctggcgcgcccgcgctcgagaacatcgccaaccgcgactactgg
tcctccgtggggcggtaccccggaatggctacctggtgctggcgcgcccgcgacattc
acgtgctcgccatcgatggagttctga

5 SEQ ID NO:8 polypeptide sequence of Orf4

MTGFHARKPVGGGHRPAHGRPLAWPRIPLRTTMMKPLPLAYLAALLPWYAGVIAQSAP
AAGDDASITLEAVRVEASADASAGGLAPAFAGGQVATGAKVGILGTRDNLETPFSITAYT
NELIQDRQAKGVGDVLDNDPGVRVARGFGNFQESYFIRGFILSSDDIAYNGLYGLLPRQY
10 ISTQLFERVEVLRGASAFITGAPPSGGGIGGVINLVKRAPNEPLTRFSAGYGSDSVLEA
SADIGRRFGPDDSVGIRINAAQRGGETAIDGERTRTTVFALGLDWRGERARLSADIGYQD
NRLKRARPNTVLGDAAKVPGAPDAGSNYAQPWSYSNERDVFGLRGEYDFNGRITGWVA
YGMRSKEENSLANLNNVNGAGQGKFYRFDNAREDTVNTGEIGLRAKARTGPGVHELVAS
ASYFDLEKKNAIVMDFNFQFDTSTYDPVSYAKPAISSTAFRGNDMDPAKQGVIRLASYA
15 LGDTMSFFDDKVLTLTAGIRHQRLYQRDYSYDTGIGGTPYEQSHNSPAAGLVVRVTPQVSL
YANYIEALSAGDTAPQTANGLPVVNHGESLAPYVSKQKEVGKFEHDLGGGLALFSTDK
PRGFVGDDQVFRASGKDRHRGVELTTYGELTRSVRVLGGLTWLDAKQLSTGNAATDGKRV
IGVPRFQANLGVWDIPGVQGLTVDGRVVYTGSSYADAANTLEVPGWTRLDAGLRYMTDI
GGHLVTWRARVENIANRDYWSSVGGYPGNLYLVGGPRTFTLSASMEF

20 SEQ ID NO:9 polynucleotide sequence of Orf5

Gtgccctgctcctacttcccgcgctacgcgcccctgcgcgcggcaggcgcagcccgccctc
Gtgcccgcgctcttcatgcttgccctggcgccgctgcccgcggcgcgcgcccagccc
Gctgcgccgggtgtcccggatacgcaaggcgtggcgcaaattgccggccgtcacgggtcaac
Gccgcgcccgtggacgacacgctggagcatctggaggcgccggtcgataccggcgcgctg
25 Ggacggcgcacccagctggagacgccttttccaccacgggtggtgactgcccgcgacatg
Gaggagcgcagggtcaacaagctgggagacgtgttcgcgctggatgcctcggtgacggac
Aacagcgcctcctatggcgcgctgggacgctacctgacgggtgcgcggcctgccgctggat
Tggcagaattcgtaccgcatcgatggccggccgcttctgagctacgtcacgacgctgccg
Ttcgagcacttcgagcagatcgacctgctcaagggggcgctgggcttcatgtacgggttc
30 Ggctcgccggggcgccctggtcaactatgtcaccaagaagccgacccgacgaagcgggtgcgc
Agcgtcgagctgggctacgtgtccaaggggttgcctgcgcgagcacgtggacctgggcggc
Agggctggccagagcggcgccgtttggcctatcggtgaatgccacgcacgaggagggaat
Acctacaacgggtggatcgctgtaccgcgattcggtgtcgctggcgctggatgcgcgctg
Agcgaccgggtgacctgggacttccaatccatctaccaggaccgcaaggccatcgggcag
35 Gagccacgatctatgcgggcacccatggccggcagcgagttgccatcgccgggtgcgcaac
Gacaatgacagggtggtcgggcagggaccgtatgcccgaacatgcgttccgctattactcg
Accggcttgaagtaccaattggcggacgaatggacgctgagcaccaattacagctacagc
Tccacgcgtaccgcggcgaacgagtcggtgctgttccgcgcgaccaggcgggcgactat
Gacgattaccgctcgactatggcgaggccatggctacaaccagtggcaggccatgctg
40 Gagggaagttcgctaccgggtcccttgaagcaccacgtggtggccggcgcgctcgaggcag
Aagcagaagaacgactacagcgccaacggggtctatcaattgcaggggcacgggcaacctg
Cgcgcgcgcaataccaacacgtactacagcgaaggccagctgcacctgtaccgcgcggcc
Gagatcacgcagaaggcgctgttcgcagcgacacggctcgacctgaccggcggtggtcg
Gtgcgtggcgggctgcgctatacgaattatgcgcagcaaggattcgatgccacggcgcg
45 Cgaacatcgcgctacgacaagaacggcgctgctgacgccgaccttgcctgatgtacaag
Ctgacggcgcgccatggcctatgccagctacatcgaatccttggagccgggctcgctcg
Gtggggcgccgctacgccaacttcggcgcatgtgctcgatccgttgaagagcaagcagtac
Gagctgggcatcaagaccgaacaggacggctgggcccacggcgcgctgtttcgcatc
Gagaagaaggcggaatacgcgaatgccgccaacgagctggtgcaggacggcaagacgctc
50 Tatcaggggttgggaactggcgccctccacgcgtatcgcccgcgactggaacgtgggagggc
Agcctgatgttgcgtggattcggaatacaagaaggctcggtttcaccggcaaccgcgtg
Gcgggagcgcggaagtctggtggcgccgcgcaactggcgctactcggtgccgcagggtgccg
Gggctgaagctgcgcgccgatgtgaagtacaccggcaacacgatgctgggcgcccagcaac
Cgggtgcagggtggacgactacgccatcgctcaatatcggcgccacctacgacacgcagatc
55 Caggctacgaggcgaccttcaccggcgcatcaacaacgtggccaacaagcgctactgg
Ctgtaccagtcgctgactacgtgaaggcgggcgaccgcggacctatggcctgacgtct
atatccagatatccggatatctggatataa

SEQ ID NO:10 polypeptide sequence of Orf5

5 VPRPTSRRTTRPARRQAPAFVPALFMLALGAVAAGARAQPAAGVDPDTQGVAQMPAVTVN
 AAPVDDTLEHLEAPVDTGALGRRTQLETPFSTTVVTARDMEERQVNKLGDVFDALDASVTD
 NSASYGAWASYLTVRGLPLDWQNSYRIDGRPFLLSYVTTLPFHEFEQIDLLKGASGFMYGF
 10 GSPGGLVNYVTKKPTDEAVRSVELGYVSKGLLREHVDLGGRVQGSGAFGYRLNATHEEGN
 TYNGGSLYRDSVSLALDARLSDLRTWDFQSIYQDRKAIGQEPTIYAGTMAGSELPSVVRN
 DNDRLVGQGPYADNAFRYYSTGLKYQLADEWTLSTNYSYSSTRTRRNESVLFRLRDQAGDY
 DDYRSDYGEAYGYNQWQAMLEGKFATGPLKHHVVAGASWQKQKNDYSANGVYQLQGTGNL
 15 RARNTNTYYSEGQLHLYRAAEITQKALFASDTVDTLGGWSVLGGLRYTNYAQGGFDATGA
 RTSRYDKNGVLTPTFALMYKLTPTMAYASYIESLEPGSSVGAAYANFGALLDPLKSKQY
 ELGIKTEQDGWAATAALFRIEKKAEYANAANELVQDGKTLYQGLELGASTRIARDWNVGG
 SLMLLDSEYKKGSDFTGNRVAGAPKFVAAAQLAYSVPQVPGLKLADVKTGNTMLGASN
 RVQVDDYAIVNIGATYDTQIHGYEATFTAGINNANKRYWLYQSSDYVKAGDPRTYGLTS
 ISRYPDIWI

SEQ ID NO:11 polynucleotide sequence of Orf6

20 Ttggcccgccataagcggtcacgggtcgtgagatttccgacctcaccgaggggtacaaacgcc
 Tacacaaccgagggccatgagcacggccacgggcctgacactctcgccacgcgaaacacca
 Caatccggtcagtggtcaccggacagcagatcgaggatcagggcctcaccgacaccggc
 Gccatcctggcgaccgcgccagggatttccggtcacgcgcagcgacagcaaccgctattca
 Ttctcgggcccgcggttcaccatcgacaacttccagtttgacggcctggatcgcccatc
 Ctgagccaatggaactatgggttcgaccgatatggacgccgccatctacgatcacgtggaa
 25 Atcggtacgtggcgccacagggcctgatgacaggctcgggcaatccttcagccgctggaac
 Ttcggtgcgaagcgctcccttgcggtgagttcgcggtacgttcaatgcgagtgctcggcagc
 Tgggactatgtgcgcggcgatgcccacatctccgtgcccatcacggaagacggcagaata
 Cgggtcacgcttggtggccgcctacagtcagggcgacagctatgtgcacttttagatacg
 Cggcgccgcacattctatggcggtgggtcagcgccgatctgacgccggatcacgggtgctgacg
 30 Accagcggtggagtagcagcacaaccacagcaatggggttggcagcggtttccgctgttc
 Tatagcgacgggttcgcgcaccgatttcaaccgctcggtggccaacaacgctccctggggcc
 Cggcaggatcacgaagccaccacctatttctggtggacctcacgcaccgcttcaccaatgac
 Tggaagctgcgcgcggcctatagccacactgatggcgctatctcatgaaacatgtgtac
 Cggggcggttaccatcgccatactggcatcatcgctgccccccctgcattttccaac
 35 Tacgacggcaacctcgatcgggatgacatccatttttcttgcctcctttcgaggcc
 Ttcggcctgcgccacgaagtggcctgggctggatgagcatcgacaaccatagcgacatc
 Cagcgatacgcaatggtcggaccggccccagccatcggcagcttcttcgactggcgccgc
 Gccacatccaagagccagctggtggcgccacagctgtcgcccgccgacgacgtgcgcacc
 Aagcagaccggcgccctatctgggtcgggcggtttgactagccgaacccctgcacctcatc
 40 Gtggggcgaccgttgagcgactggaaaaccaaacagatgtattttgggtcgcgcccgcaa
 Tacaggatcaagaatcagttcacccttatgcccgtctgacctacgacatcaacgacacc
 Tacacggcgtacgccagctatacggagatcttcagccgcagaacgcgcgcgacaccagc
 Ggcggcattcttctccatcaaaagcaagagctatgagctgggtctgaaggcagcctat
 Ctgaggggagcgtcaataacctccgcgcgctctttcagacgcggcaggataacctggcc
 45 Caggtcatcccggtcatccattccgggctttccgaacatgcaggcctcacgtgcccgc
 Tccggcgccaaggtcgaggggatagacctggaggccagcgccagatcctgcccgactgg
 Aacatcggcgcagctatacacacttcaccaccaaggacgccagcggaaccccatcaac
 Accaatcatccgcgcagcctgttcaagctctacaccacgtaccgctgcccggcgccctg
 Caccggcttaccgtggcgccggcggttgactggcgaagtgcagctacgcactcgtgagcctcatggcg
 50 Agtcgcgcggcaatgtcgaaagtcgaacaggacagctacgcactcgtgagcctcatggcg
 Cgcttcgactttaacaaaaactgtcggaacactgaacgtgaacaatctgttcgacaaa
 Aagtactacgatcagatcggttctacagccagggttggtgggggtgcgccacgcaatgta
 atgctcaacttgcggggcgagattga

SEQ ID NO:12 polypeptide sequence of Orf6

55 LPAISVTGREISDLTEGTNAYTTEAMSTATGLTSLPRETPQSVSVVTRQQIEDQGLTDTG
 AILATAPGISVTRSDSNRYSFSARGFTIDNFQFDGLVSPILSQWNYGSTDMDAAIYDHVE
 IVRGATGLMTGSGNPSAAVNFRKRPLREFAAFTNASVGSWDYVRGDADISVPITEDGRI

RSRLVAAYSQGDSYVHFLDTRRRRTFYGVVSADLTPDVLTTTSVEYQHNHNSNGFSGSFPLF
 YSDGSRSTDFNRSVANNAPWARQDTEATTYFVDLTHRFNDWKLRAAYSHTDGRYLMKHVY
 RGGYPDRHTGIIAAPPAFSNYDGNLDRDDIHFSLSAPFEAFGLRHEVALGWMSIDNHSDI
 QRYAMVGPAPAIGSFDDWRRRAHIQEPSWADTLSPADDVRTKQTGAYLVGRFALAEPLHLI
 5 VGDRWSDWKTKQMYFGSRREYRIKNQFTPYAGLTYDINDTYTAYASYTEIFQPQNARDTS
 GGILPPIKSKSYELGLKAAYLEGRNLNTSAALFQTRQDNLAQVIPGSSIPGFPMQASRAA
 SGAKVEGIDLEASQILPDWNIGASYTHFTTKDASGNPINTNHPRSLFKLYTTYRLPGAL
 HRLTVGGGVWDQSRMYQAAASPRGNVEVEQDSYALVSLMARFDFNKKLSATLNVNNLFDK
 KYDQIGFYSQGWGWAPRNVMLNLRAQY

SEQ ID NO:13 polynucleotide sequence of Orf7

Atgaagttctactcttcccatccgatgcccagctcgctcgcggtcgcatcgagtcgct
 Ctgttgggctgctgcccggcgccagggccgctccacggcggtccagctgccatccgctc
 Acggtcgagggcgagtagctcgctcctatcaaccggaaagcgcccagtcgcccaggttcacc
 15 Gcgccccctggcggacacgcccgcacgggtcgaggtcatccctgagcgggtcatccaggac
 Cagggggccagatcaccctcgaagcggtagctcgcaatcgccagggatatacgatgaccgcc
 Gcggaaggcgccgctccggccagcgacctgcccgttcatccgcccagaaattcgccagc
 Agcctttttgtcgacggcctgcccgatcccagcacgcaatcgccgcatacctcaacctg
 Gaacaggtcgagctcgtaagggggccgattcggtattttccgggcgcccggcgccggc
 20 Ggaagcatcaacctcgtaaccaagacgcccaggaaccaggtttcaccgaagtcaggcc
 Ggcacgggacggccgagacctaccgaggcaccatagacggcaactgggtgctgggagag
 Aacacggcgctgcccctcaacctgctgggcaccagggacaccgtgcccggccgcccgaag
 Gcggtcgaggttcagccgctgggtatcgcccatcgctgcccctgggctgagcggcccc
 Acccgcgtagcgtgggctgtaccactatcgccaccggcggttcccgatattcgatt
 25 Ccgtagcatccgcccacggcgacgcccgatcaccgagaccatcgcgctcagccgcccgaac
 Ttctacggcctgggtgcccgcgactccggcgataccgaggactacgcccaccgctcaa
 Tgggagcacgacctcgcaatggcttcaaggtggagaacctggcgcgctactcgcgctgcc
 Acggtggagcagatcaccacatgcccgaactgaaacccgcccgatctggccaaggggctg
 30 Gtgtagcgaatctgcccgcagctaccaggtcaacgacagtttcgccaaccgacccgac
 Ctgcgcggtacattcgacacggggcagtgccgcataccttcgatctgggcggggagttc
 Gccaccagccggcgagtcgcccagcgtacaagcaagaaatccccgacgcccggcagtcct
 Tgctgcccgtgacggacggcaacaatccccgcccgtgcccgcctcgctccgggatccggat
 Ccgacgtgggatttcccgggaacgggtgcccgcgcaaccataaccggcccgtaccacacc
 35 Gacatcctgtccctgtacggtttcgacaccatcgccctcgacgagcagtgccagctgaat
 Ctggcctgcccgtgggaccactacaagaccagcggaacacctgcccgtacgagggcgcc
 Aagccgcccgtctacgagcgtgcccgcgcccagcacaacctgttcaactaccagctcggc
 Ctgggtctacaagcctcgctccggacgggtcggtgtatgagagttacggcacggcgctccacg
 Ccgtagggcgtgctccgactacgccccggcggaacagcatctccggcacaagccagcagctc
 40 Aagccggagcgcagcagggcgatcgagatcgggacaaagtgccaggtgctggaccggcg
 Ctgtggtgagcggcgccatgttccgagacgcccgaagaacaccagcatcgaagtcgccc
 Gaaggcctgcccgcaccagccggaagagccgctcaccggcatggagctgggctggcg
 Ggcagcctgacgcccgcgctgggacgtctacggcggtacgcccgtgctcgacagcaagctg
 Gtcaggccagccataagagcggggcgcaaggccagccgctgcccagcgcgccccggcac
 45 Gcattcagcatctggagcacctacaagctgctgcccgaactgaccgtgggggcccggcg
 Ttctatcgagcaaggtctatggcaacgcagatgcccgtatacaacaaggacggcacgccc
 Aaggcgcgctgggtgcccgcgtactggcgcttcgacgcccagggcggtaccagcttaac
 Aagcaccttacggcccagttgaacgtctacaacctgctcgacaagacctattacgccaag
 Acctaccgagccattacgcccgcgtggggccggggcggtccgcatgctgacgttcaag
 ctgagctactga

SEQ ID NO:14 polypeptide sequence of Orf7

MKFYSSHPMPESLAAAIAPVLLGLLPAAQAASTAVQLPSVTVEGEYSSYQPESAQSPKFT
 APLADTPRTVQVIPERLIQDQASDLEAVLRNAPGISMTAGEGGRPASDLPFIRGQNSAS
 SLFVDGLRDPSTQSRDTFNLEQVDVVKGPDSVFSGRGGAGGSINLVTKTPRNQDFTEVQA
 55 GIGTAETRYRGITDGNWVLGENTALRLNLLGTRDTPVGRDKAVEFSRVGIAPSLRLGLSGP
 TRVTLGLYHRRHRRVPDYSIPYDPRGTGTPITETIGVSRNRFYGLVRRDSGDTEDYAATVK
 WEHDLANGFKVENLARYSRATVEQITTMPELKTADLAKGLVYRNLRASYQVNDSEANRTD

LRGTFDTGQWRHTFDLGGFATSRRSRDRYKQEI PDAASPCSPVTDGNNPALCASLRDPD
 PHVDFPGTVRRNHNPNARYHTDILSLYGFDTIAFDEQWQLNLGLRWDHYKTSGRNLPVRGA
 KPPVYERAARTDNLFNQYQLGLVYKPRPDGSGVYASYGTASTPSAVSDYAPADSI SGT SQQL
 KPERSEAIEIGTKWQVLDRLRLVTGAMFRETRKNTSIEVAEGLRAPAGKSRVTGMELGVA
 5 GSLTPRWDVYGGYALLDSKLVRAHKSQAQGPLPSAPRHAFSIWSTYKLLPELTVGAGA
 FYRSKVYGNADAGYNKDGTPKARWVPAYWRFDAMAAYQLNKHLLTAQLNVYNLLDKTYAK
 TYRSHYAALGPGRSAMLTFLKLSY

SEQ ID NO:15 polynucleotide sequence of Orf8

10 Atggagaagccgttgaaatccctggactcgtattcagcgagcacgctcgccaactcgctg
 Gccgccgccattgcggtgcccggccctgtgctgatgcccggtgctcaggcacagaccagc
 Gcggcggttaccgaattggcgccgggtgcaggttagagggcgaagcgtccccctatcaggcc
 Accaccgtccagtcgtccaagatgacggcgcccttgctggatacggccaggaccgtgcag
 Gtcgtgcccgcagcaggtcatccaggaccggccggccaccaatctgcaggacgtgctgcgc
 15 Aactcgccgggcatcaccatggggcgccggcgagggcgggcgccggcgccgacgtgccc
 Atcatccggggccagaatgcccggcgagcatcttcgtcgacggcggtgcgcgacccagc
 Acccagatacgcgatacgttcaacctggagcaggtcgagatcatcaaggggctgattcg
 Gtctactccggcgccggcgagccggcgccgagcatcaacctggtcagcaagacgccaag
 Gcgcgcgacttcgcccagggctcggtgcagatcgccaccgacagcaattaccgcccacc
 20 Gccgacggcaactggctgctggcgcaacgcgccttcgcctgaacctgatgggcaac
 Aagggcgagctgcccggcgccgaccatgcggctgatttcagccgctggggcggtggcgccc
 Accctgcaactggcgctgggcacgcccaccgcgcatcaccctgggggtactaccactaccag
 Gatgacagcatgcccgttaccgcatcccgtagcatccgaagtggggcgagccggtcacc
 Gagaccagggcgctcagccgcaagaattttctacgggctgaccggccgagcttcatgaag
 25 Tcgcgcgacgagtggtccacgctggccatcgatcacgatttcagcagcaagctgcgcctg
 Cgcaacgtcaccgctacggcgctcggtgaccgactacgcccaccacaatccggatgac
 Agcaagggcaactgcccacgggctgggtgtaccggcgctgaaggcggtactacacc
 Aacaagacgttcaccaaccagaccgacctgagcggcgcaattcgagacgggcagcctgcag
 Cactcgcttcgacgtgggcttcgagtagcaacatcaagcaggacaaggactcgtatacc
 30 Cagactatcgccaagggcgcgatgccttgcaaggtggggcgccaacgatgccagcaatccg
 Gccttggtgcacctcgctgtgggatccgctatccaccgacacgcatcgcgctctacggcttcgacacg
 Cgcaacgacaaccggcccgctattccaccgacacgcatcgcgctctacggcttcgacacg
 Atcaagttcaacgagcaatggcaggccagcgtcggtgactgcttgggacaattaccgcgta
 Agcggcgagcaatatcgcccgccggcgcaacgatcccgccagcagcggcggttctacagc
 35 Accagcccggaagacaatctgttcaattaccagctgggctggcctacaagccgggtgccc
 Aacggcgagcatctacgcctcgtatggcacctcgctcgacgcgctcgcccgctcgccggtcg
 Aacgtgagcgagccggtgacggtgagcaacgagtcgctggcgccggagaaaagccgcacc
 Gtcgaggtcgccaccaaagtggcaattgttcgacgacgcgctgacctgtcgggcgcgcttg
 Ttccaggacatccgcaagaacaccagcgtggcgctgtcgggcgaccgaaacgggagcagatc
 40 Ggcaaggccaaggtgcgcggcatcgaaactgggcttctcgggcgacatcacgcccgaagtgg
 Aacgtctacggcggtataccttcatggacagcgaaactggtcgagggcgccctacaacagc
 Ggcgcggtgggcccaggacctgcccacacgcgcgcaatgccttcagcctgtggaccacc
 Tacaagctgggtgcctcagctgaccgtggggcgccgctattacgtggacaaggtatat
 Ggcaacgcggacaacggctcgcaatgccgacggcgacggcggaaggcgcgctgggtaccgtcg
 45 Tactggcgcttcgacgcatggccgctgacgagttcaacgaccacctgaccgcgagctc
 Aacgtgatgaacatcttcgacaagacgtactacaccaaggcctacgcccgcactacgcg
 gcgctggggcacggggcgccgcggtgctgtcgttcaatatcaagtattga

SEQ ID NO:16 polypeptide sequence of Orf8

50 MEKPLKSLDSYSASTLANSLAAAIAVPALCLMPGAQAQTSAGVTQLAPVQVEGEASPYQA
 TTVQSSKMTAPLLDTPRTVQVVPQQVIQDQAATNLQDVLNRNSPGITMGAGEGGRAGDLP
 IIRGQNAAGSIFVDGVRDPSTQIRDTFNLEQVEIIKGPDSVYSGRGGAGGSINLVSKTPK
 ARDFAEGSVQIGTDSNYRATADGNWLLGDNAAFRLNLMGNKGDVPGRDHAVDFSRWGVAP
 TLQLGVGTPTRITLGYHYQDDSMPPDYAIPYDPKSGQPVTTETQGVSRKNFYGLTGRDFMK
 55 SRDDVATLAIDHDFSSKLRLRNVTYGRSVTDYAATNPDDSKGNVPNGLVYRALKAGYYT
 NKTFTNQTDLSGEFETGSLQHSFVDFEYSNFKQDKSYTQTIAKGAMPCKVGDANDASNP
 ALCTSLWDPDPHDYYPGHLNRNDNPNARYSTDTIALYGFDTIKFNEQWQASVGLRWDNYRV
 SGNRIARGNDPASTPAFYSTSREDNLFNYQLGLAYKVPVNGTIYASYGTSSTPSAVAGS

NVSDAVTVSNESLAPEKSRTVEVGTKWQLFDDRLTLGALFQDIRKNTSVAVSATETEQI
GKAKVRGIELGFGSGSITPKWNVYGGYTFMDSSELVEGAYNSGAVGQDLNTPRPAFLWTT
YKLVPLTVGGGAYYVDKVGYNADNGRNADGTPKARWVPSYWRFDAMAAAYEFNDHLTAQL
NVMNIFDKTYTAKAYAAHYAALGTGRAAVLSFNIKY

5

SEQ ID NO:17 polynucleotide sequence of Orf9

Ttgcatactcgacgccacagcgccggtcgcgccgacctgctgcatctatcg
Ctggccgacctgctggcgccggtcgccagcgccagccaccgaagccaccacc
Ctggccaccgtgcaggtcaccggcagggcgagaccgccaccggcccggtcgacggctat
10 gccgccacgcgagcgccaccgccaccaagaccgatacccccgtgtcggaacccccgcag
gccgtcacggtgatcccgcgcgagcagatcatcgaccagggcgcgagaaacgtgcaggac
accatgaactacggcggggtgcccccaacgcctatggcggtggacaaccgcgggcag
tacgtgcgggtgcgcggggtggagccggtccagatctcgacggcctgaagcagttcttc
aactacaacaatcccgcgaccgaggtctacgggctcgagcgctcgaagtcccgcgggc
15 ccggcctcgatgctgtacggccagggcagcaccggcgcggtggtcaacctcgtagcaag
cggccgagccccgagggccatgcgcgagatcgcggtgaccgtgggcaaccacaaccgcaag
gagatccaggccgatctcaccggccccgtgaccgaggacggcacctggctgtaccaggtg
gtggccctcgagcgcgacagcgacgcaggtccagtacaccaaggacgaccgcatgatg
ctcgtgcgctcgctgacctggcagccgacggccacctcgctgacgctgcaggcctac
20 tggcagaaggacaagtccggcgacgaccaggccttcctgccctggagcggcacggtcagc
ggcaacccccaacggccgatccccaccgcccgttcaccagcgaacccggcttcgaccgc
tacgacaccgagcaattcagcgtgggctggcagttcgagcacaagttcaacgacaactgg
aaagtgcgccagaacctgcgccacacatccagcaaggtcgactacagcagctgtatccg
gcggtctacggcaacccccgacaatcccttcacgacggccgaccagcgctggtcaatcgc
25 tacctgtacatcaagaaccccgcgcatgcttccttgctggccgaccagaacctcgaaggc
aaggtgaactggggccgcccgaacataccctgctgatggcggtggactacagccgctat
cgcgagaccggcgagaccggcagcggttcggcgcgccgctggacctgtaccagccggctc
tacggcaccctgcccagctatgcatgtcgagcgtgcccaagaacaagcagcagcagatc
ggcgtctacctgcaggaccagatcaagttcgaccgcaactggatcggtggcgggcctg
30 cgccacgaccgctcgccaacagcgtcgagggcgccgacaaggaaaccgacaacgccacc
accaagcggctgggctgatgtacgcgcccgaacaacggctggctcgccctacctcagctac
agcgagtcgttcacccccatcgccggcaccgacaacagcggaacccgctgggtgcccgatg
cgcggaagcaatgggaagcagccctgaagtacatgccgcaggacaccggctatgaggcc
acctggcgccctacgacctgcgcgagcgcaaccgcccagaccaacgaccgctccgatccc
35 accaaccaggtgcagaccggcaagaccaagacgcgcggcatcgaactggaattccgcggc
cgctcagcgccgagatggatgtgatcgccaaactacaactacaccgacatcgaccggcag
ctcgaaggcctgcccagcacacgttctcgctgtggagcaaataccggttcagcggtgggc
gatgtgcatggcttgcgcggcgccggcgctgctacctgaacgcgttctcgcgacggg
tccgcgcccagaccggctcggtggccctgttcgacgccatgctcagctacgacaccggg
40 tcgtggcgctatgcgctgaacgtcgccaacatcgccgacaagacctacgaggtgggtg
ctgcggcgccgagattgcttctacggccagcgccgacgggtcacctgagcgccatgtac
cgcttctag

SEQ ID NO:18 polypeptide sequence of Orf9

45 LHTRTPQRQRFVAPRLLHLSLAASLAAGAAQAQTATEATTLPTVQVTGRGETATGPVDGY
AATRSATATKTDTPLETPQAVTVIPREQIIDQGAQNVQDTMNYAAGVRPNAYGVDNRGD
YVRVRGVEPVQYLDGLKQFFNYNNPRTEVYGLERVEVLRGPASMLYQGSGTGGVVNLVSK
RPQPEAMREIGVTVGHNHRKEIQADLTGPLETGDGTWLYQVVALGRSDTQVQYTKDRMM
LAPSLTWQPSAATSLTLQAYWQDKSGTTQAFLPWSGTVSGNPNGRIPTRRFTSEPGFDR
50 YDTEQFSVGWQFEHFKFNDNWKVRQNLRLHTSSKVDYSTLYPAVYGNPDNPFIDADQRVVNR
YLYIKNPRMRSLLADQNLLEGKVNWGRAEHTLLMGVDYSRYRETGETGSGFGAPLDLYQP
YGTLPDYAMSDVPKNKQQQIGVYLQDQIKFDRNWIVVAGLRHDRVANSVEGADKETDNAT
TKRLGLMYAADNGWSPYLSYSESFTPIAGTDNSGNRWVPMRGKQWEAGLYKMPQDTGYEA
TLAAYDLRERNRQNDPSDPTNQVQTGKTKRGIELEFRGRVTPQMDVIANYNYTDIDPQ
55 LEGLPKHTFSLWSKYRFSVGDVHGFAAGAGVRYLNAFRDGSAPETGSVALFDAMLSYDTG
SWRYALNVANIADKTYEVVCLRRGDCFYGQRRRTVTLASMYRF

SEQ ID NO:19 polynucleotide sequence of Orf10

Atgggtttacgcttgcatcagggcgcccgccgccccgtcccgcgccgccaaggcgcccc
 Caaacggcactggccatgcgcgccgctggcgccatgcgactggccggtacgctggcg
 Gccgctcccgcgcccgcgagccgacggcgcccgcccgcatcgcgggcgcgccgctgg
 caccatcgacgcccggccccctgggagggccctggcgcgctttgccgaccaggccggcatt
 5 accctgctgtacgaccccgccgcggtgcgcgcccgccagcgccggcctgcaaggcggtg
 tactcggtgcccagcgccctggcgcgccctgctcgatggcagcgccctggacgcgccag
 cgcgcgccggcacctacgtgctgacggcgctgcccgcgcccgggtcgcccagctggcg
 ccggtcaccatcgaggctgacggcgctgcccgcgccgatcccgcctgggcccgcaccgccaag
 cgccgagctcgacgcgccgaggtgctcgactggagcgatatcggaagcgcgctcgat
 10 cccggcgtaactacaaccgcccacccaagagcatcaacatccgcggcctggacgaaaaac
 cgcggtggtcacgcgcatcgacggcatccgcctgcccctggctcgacgacggcgcgccggc
 atccagggcgggctgaacgcggtggacttcaacaccctgtcgcgccctggacgtcggtgcg
 ggcgccgactccagcgcgccggctccggcgcgctggcgccgctggcgacctgcgacg
 ctggaacccgcccagctgctgcgcgacggcgccgcttggcgcgctggccaagtccgac
 15 tatgactcgcccgacgcccagctggggcctgaacgcgccctggcgggcgaggtccacgac
 gacaccagctgggtggtgacggcgccacccgcaatgggcacgacctggacaaccgcgcc
 gacacggcgccgtacggcagcaagcgacgacccagccccgaggactacgcccagaac
 aacttctgtctcaagctgacgagcgcatcgacggcgcccatcgcccgccctgacgggc
 gaatacttcaagcgccgcccagctcgacagatgtaccagcagggcgcccgccaccagc
 20 taccagtacggcgccaaccgacccacgaggaaccacgcgcaagcgcgctctcgctggac
 taccagtacaacgcccgcaggccggcgccgcatcgacagcgcccgcccgccatggtgtat
 tggcagcggtgcggtggacagctcgacggcgcccgcccgacgcgcgacggcgccg
 tacgcccggcgccgacccgtacttctacggctaccccagcgcccccctacggcgccgacg
 aactcgatccaggaatcgatcctcgggcgtaacggcgagctctccagccgcttcaaggc
 25 atggtgtcgacgcgctgacgataggcgccgaatggtacggcaaccgacccgagcagtac
 tcggacggctacgacaactgccccgcatcccgcccgccacgcccgcgccgatggggcg
 cgccgtgctgacatgctgcataccaaccaggccgacatgccccgggtcaaggcgccag
 tgggcatctggggcgaggacgaaatcgccctcgccgacggcgctacatcctgacccccg
 tcaactgctgacgaccattacgagcagaagccgacgaaggcgccggtaccagaacaac
 30 cccaacgcccggcgcgctgcccgcgtcgctcgctggggggcgcccttctcgcccaagctgctg
 ggcacctggaaggcgcgcgagcgctgacgctgtatgcgcaatacggcttcggctaccgg
 ggcgcgtcgccaccgagctgtacaccaactacggcgcccgccggaacctatctgcgcgctg
 ggcaatccctccttgaagcccagaccagcaagggtgggaactggcgcccgccctgggc
 gacgaccagtgggagggcgccgtatcgctgttcgacaaccgctaccagaacttcatcgac
 35 aagaacgtgcccgtgggcaagggttcgcccgaatggcagccggcctgggacggccagtac
 ccgctggcgctcaccgggctggccaaccggcgcgcgctgcgcatctatggcgccgaagcc
 tcggcgccactggcggttcgcccccaactggcgccctggggctcgctggcctgggcccgtg
 ggcaaggacgaaaacacccggccagcactgaattcggtgcccgcgctcaaggccatcctc
 ggcctgggctaccagcgcgacgaatggggcatcgacgcccgtgacggccgcccagcgcg
 40 cgcgacgagctgcaataccccgagggcctccgcccagcgcgctacgcccatttccagggc
 ccgggctacggcggtggatctgtccgcctactggcgcccggcgccgtcaagggcctg
 cagctgcaggcgggcggtgttcaacctgttcgacaagaaactgggaagccatcaacgtg
 cccacggcggggtgccattgcatcccgcgaccgttagactggtacaacgagccaggccgc
 45 agcgtgcgcgatcggtgacctaaccagtattga

SEQ ID NO:20 polypeptide sequence of Orf10

M V Y A C D Q G A R R A P P P R P Q T A L A M R G A L A A C A L A G T L A A A P A A A Q P T A A P A S A G A R A W
 H I D A G P L G E A L A R F A D Q A G I T L L Y D P A A V R G R A S A G L Q G V Y S V P D G L A R L L D G S G L D A R Q
 R G A G T Y V L Q A L P A G P V A Q L A P V T I E A D G V R A D P A W A R T A T R R E L D A R Q V L D W S D I G K R V D
 50 P G V N Y N R R T K S I N I R G L D E N R V V T R I D G I R L P W L D D G A R G I Q G G L N A V D F N T L S R L D V V R
 G A D S S A A G S G A L G G L A D L R T L E P A D L L R D G R R F G A L A K S D Y D S A D A S W G L N A A L A G Q V H D
 D T S W L L Q A G T R N G H D L D N R A D T G G Y G S K R S Q P S P E D Y A Q N N F L L K L Q Q R I D G G H R L G L T G
 E Y F K R R A D L D Q M Y Q Q G A G T S Y Q Y G A N R T H E E T T R K R V S L D Y Q Y N A P Q A G A A I D S A R A M V Y
 W Q R L R L D S S Q D A R R T R D G R A Y A R P G D P Y F Y G Y P S G P Y G R S N S I Q E S I L G V N G E L S S R F E G
 55 M V S Q R V T I G G E W Y G N R T E Q Y S D G Y D N C P A I P P G T P A P M G P R L C D M L H T N Q A D M P R V K G S Q
 W A I W A Q D E I A F A D G R Y I L T P S L R Y D H Y B Q K P Q Q G G Y Q N N P N A G A L P P S S G G R F S P K L L
 G T W K A R E A L T L Y A Q Y G F Y R A P S A T E L Y T N Y G G P G T Y L R V G N P S L K P E T S K G W E L G A R L G
 D D Q L G G A V S L F D N R Y Q N F I D K N V P L G K G S P Q W Q P A W D G Q Y P L G V T G L A N R A R V R I Y G A E A
 S A H W R F A P N W R T W G S L A W A V G K D E N T G Q H L N S V P P L K A I L G L G Y Q R D E W G I D A M L T A A T R

SEQ ID NO:21 polynucleotide sequence of Orf11

45 **SEQ ID NO:22** polypeptide sequence of Orf11

111

SEQ ID NO:23 polynucleotide sequence of Orf12

5 Atgaaatcccgtcactccggcgctgcgccgggtgtcctggcctgtgtcgctccgttggcc
 Ggccacgcccaggccggcgccgcccggcccaacccatccccgaactcgatccggtcgtc
 Gtcaccgcccgcgcatcgccccagctgctcaagaatgtgctggccgacgccagcgtgatc
 gagcgcgatacgtggcgcgccggccagtcacgacctggccgaagtgtggcgagcag
 cacggcatcgaattcgccgacagcgggcccgcaaagcgtcaccagcctgttcatgcgc
 ggcgccaacagcaaccagaccctggctcgtcaacggccagcgcatacaaacgccaac
 10 ggcgccggcatttgcgctcaacgcgctgcccgggaagccatcgaacgcatacgagatcatg
 cgtggcgccggccagcagcctgtacggggccgacgcgacgcccggcgatcaacatcatt
 acccgcgagccggcgacaaggcgctgtcggcctatgccaacgccggttacggcacctac
 ggcaccagccgctacgacgcggcgctcctggcgccggccgacggcttcagctacagcctg
 tccaccggctatggccagagccatggctcaacgccaccaacccgctcgttctcgtaac
 aaccggacaaggacagctactaccagaactacgccaccggccagcgtgggctacgaatgg
 15 cggcccgagcagaaagtgggtggcgaggtctaccgcagccgcataacggcggttacgac
 gcctcggcctcgtacgactacaacgaccgctacatccaggacctgcaggcctattcgtg
 gccagcgaacaccgctgacccgctactggaagagcagcgtgcgcgcccgtatgtggaa
 gacaagaacgattcgcgccggaaggcatgttcgaagacaacaacacgcgcttccggacc
 cgccagatgcagctacgtgtggcagaacgacttaccctggccgcccggccagcgtgacg
 20 ctggcctacgagcacctggaccagcgccgacggccagatgagcaccgccaccggcatc
 ggcaactacaccgagacgcgcggccacgtgaactcgtacaccggcgcttacctggcgcat
 ttcggccgccaccatgtgcaggccagcctgcgcaacgacaacactcgagttcggcagc
 cacaccaccggcgccctggcctacgggttcgacctgacgcccacactgcgcgccaccgtg
 gccgccaacacgggcttccggcgccgctcgttcaacgatctgtacacgcccagcagcg
 25 ttcggctatcgccggaaccccgacctcaagccgggaagagtcgcgcaacgccgagatcgcc
 ctgaaataccaggacgaggacagcgaactggcggtgtattaccagaccgcgcatcaag
 aacctgatccaggtgaccgaggacttcagcacggctcgagaacgtcgggcgcccccgcctg
 caaggcttcaccatcagcggcgccgaccgcttcgggcaacacgcgcctgcgcgccagcctg
 gacctgagcaaccgcgcaacgaagacaccggcaagcaattgctgcgcgcgcccgcacg
 30 gtgctgcgcgcccgcacatcgacctcgcttcgaccgctgctgggtggcgccgagtggtac
 gcctcggacgcgctacgactacggcttccccgaggaaaagcgctggcgggctacggc
 ctgggtcaacctgaccgcggcctacgacctgagccgcaacatgcaggtgcaggtgcgctgg
 aacaacgtgctcgccagcgctacaccttggccgacggctacaacacggccggctcgaac
 35 gccttcgtcaaccgctcgtggcgcatgtag

SEQ ID NO:24 polypeptide sequence of Orf12

40 MKSRSLRRCAGVLACVAPLAGHAQAGAAAGQPIPELDPVVVTAARSPQLLKNVLADASVI
 ERDTLARAGQSSLAEVLAQQHGEFADSGGPQSVTSLFMRGANSNQTLVLLNGQRINNAN
 GGGIALNALPPEAIERIEIMRGAASSLYGADAIGGVINIITREPGDKALSAYANAGYGT
 GTSRYDAGVSGAADGFSYSLSTGYGQSHGFNATNRRSFSYNPDKDSYYQNYATGTLGYEW
 RPEQKVVAQVYRSRINGGYDASASYDYNDRYIQDLQAYSLASENRLTRYWKSTLRAGYVE
 DKNDSRAEGMFEDNNTFRTRQMQLWQNDFTLAAGQTLTLAYEHLDRADGQMSTATGI
 GNYTETRRHVNSYTGVLGDFGRHHVQASLRNDNNSQFGSHTTGGLAYGFDLTPNLRATV
 AANTGFRAPSFNDLYTPTSAFGYRGPNPDLKPEESRNAEIGLKYQDEDSSELGVVYYQTRIK
 45 NLIQVTEDFSTVENVGRARLQGFITISGAHRFGNTRLRASLDLSNPRNEDTGKQLLRART
 VLRAGIDHRFDRLLVGAEWYASDERYDGFPEEKRLGGYGLVNLTAAYDLNRNMQVQVRW
 NNVLGQRYTLADGYNTAGSNAFVNPSWRM

SEQ ID NO:25 polynucleotide sequence of Orf13

50 Atgattccaccttgcgcgttatccctgatcccgccgctggccgcatggcgctggcagggc
 Gcctttcccgccgagcggggccgcccggctgaattggcgcccatcgcggtcatcgcc
 Gacgatcccgacgatccgcccgtattcgaaggcagcaccgccaccgctaccgccacaccg
 ctgcccggaggtgcccgcagacggctcgacaccgtgaagggtgccggacgcccctgaactatggc
 gcgcgcacgctggcgagggcgctggccggcgctgcccaatgtcaccgacgccagcgatacc
 55 cgcttcgacggcttgcgcatacggcggttcgacgcccggcagcgacttctacctggacggg
 gtgcgcgatgacagccagtagctgcgcgacctgcacaacatcgagcgcatacgaggtgctc
 aaggggcccggccggcgttctgtacggccggcagccagggcgccatcgtcaatccgggtg

5 agcaaggcgccccgggcccgggcccgcgcttccaccctcgaagtccggctgggcccgcgaggac
 tttcgcagcctgtacgccgacctgagcgcggacccttccgacacggtcagcctgcgcctg
 aacgtgggcccgcgagaatgcgggcagtttcaggcacgggtcagctcgcgcgccgcgctg
 gcgtcgccccgcttggcgtggcgcatctacgccacggctcgattggctggcgagtacgaa
 10 cacagccgctacgaccgctgcccgaccgcggcattccctcgggtggacggccggcccgcg
 ccggtcgggcccgtcgaccgtctacggcgacccccgggcccgcgacaatatcgacgatcgggtc
 cagggtgctgcgctcgcgcctgcgctaccgggcccggccaatggatgggagctgcgccatacc
 ctgtcgacgttccggctgcatagcgatttcgacaacacctatctgtccggctggcgcgcc
 gagaccgggctgggtgcaacgccagcgctggcagcagcacctgcgcgcccgcatctttac
 15 aacgtcttcgagggcccagggcacgttcgccaccggctggctcgaacaccgcttgctggcc
 ggctcgagctgggagccagcatcgcatccgacgctgcaccgcgcggccaccaaggc
 cccggcgcgagccggtgcccgggctggcgctgcaccaccccgacttgagccagcagcac
 cacggccgcatggagcgcgcccagcgatgcgcgtcaccgcgtgctgacgaaggctactac
 ttgcaggatcaactgcgattgagcagctcctggcaggtgggtggcgggcccgcgcgctggac
 20 cggttcggggtgcgcacgcgcaatcgctgctgggctggaaggcagccgtggcgaccgc
 agtgtagtccgcgcctgggagtggtctggacgcctggcggcgccagcggttctacgcg
 tcgtacagcaagactttctcgccaccggcgggcaccataggcatcacgcgggacgcg
 cggggcaacgccaatgatctgccgcccgaacatacgccagtagcaagccggggtcaag
 agcgactggctggacgggcccgtgagcaccatgctggcgtctaccagctcgaactctac
 25 aaccgcccgcgcgcgccccacgatcccacgcggatactcctgacgggcccgtgcagcgc
 tcgcgcggcctggaaatgagcggggcgggcggttagctgtgaagattcaatag

SEQ ID NO:26 polypeptide sequence of Orf13

25 MIPPCRLSLIPALAMALAGAFPAPSGAAPAEIPIAVIGDDPDDPRVFEGSTATRTATP
 LREVPQTVDTVKVPDALNYGARTLGEALAGVNPVTDASDTRFDGLRIRGFDAGSDFYLDG
 VRDDSQYVRDLHNIERIEVLKGPAGVLYGRGSQGGIVNRVSKAPGPGRASTLEVRLGGED
 FRSLYADLSADPSDTVSLRLNVGGENAGSFRHGVSSRRRLASPALAWRITPRLDWLAQYE
 HSRYDRVPDRGIPSVDGRLPAPVGRSTVYGDPRDNIDDRVQVLRSLRYRAANGWELRHT
 30 LSTFRLHSDFDNTYLSGWRAETGLVQRQRWQHLLRARHLYNVFEAGTFATGWLEHRLLA
 GVELGSQHRDPTLHRAATKGPAGQVPVPLALHHPDLSQQHHGRMERASDARHRVRTQGY
 LQDQLRLSESQVVGARLDRFGVTRNRLLGLESGRGRSVSPRLGVVWTPWPAHAFYA
 SYSKTFSPGGGTIGITPDARGNANDLPPEHTRQYEAGVKS DWLDGRLSTMLAVYQLELY
 NRRTRAPHDPTRILLTGLQRSRGLEMSGAGRLAVKIQ

35 SEQ ID NO:27 polynucleotide sequence of Orf14

Atgaacacgctgcgacgcctgcgcattcctgggcccgcgcgccacgctgggcccgcggcc
 Gccgcgcaggaggcgcccgccatgctggagccggtgcgcattcagcggcacgcgcaccggc
 Acctcgggtgctcgatacggccgctcgtggacgtggtcgatggccacgagctgcgcgcg
 40 Cgcaacctgcaggtcaacctgtccgaaggcttggccggcgtgcccggactgcagctgcag
 Aaccgccagaattacgcgcaggacctgcagctgtcgatacgcggcttcggcgcgctcg
 Accttcggcgctgcgcggcgctgacgtgacgtggacggcatcccggccaccatgcccgc
 Ggccaggccagacctcgaacatcgacatcggtcggccggcgcgctggaagtgtgcgc
 Ggcccgttctcgccctgtacggcaattcgtcggcgggcgtggtgcaggtgttcaccgaa
 45 Cagggcagcgatccgcccagggcgacgggcagcgcggcgggcggcagcttcggcacctgg
 Cgctacggcgccaagctgcgcggcgccagcgcggcagacggcctggattacgtgctggac
 Ttcaatcgcttcacgaccgagggctatcgcgaccacagcgcgcgcgcaagaacctgggc
 Aacgcgcggctggcctgcgcattggacgacggcagcgcctgacgctgagcgccaaccac
 Gtggacctgaccgcgaggatccgctgggctgacgcgcgagcaattcgaggacgaccgg
 50 Cgcagcgcgcgggtggccgagcgttcgatacgcgcaagacgctgcccagacccagggc
 Ggcctgctgtacgagcgcgcttcgacacgcgcaacgacctgcgcgtgatgctgtactac
 Ggacaacgccgaccacgcaataccaatccatccggtggcgtgcagcaagccccacg
 Caggccggcgcgctgacgacctggcgcgactacggcgcgccgacctaagctggacc
 Tcgcgccagcaggtggcggcctgcccgtgacctgatcgccggactggcctatgacacc
 55 Atgaaggagcagcgcgcggctacgacaactacaccggcccgcgcctgcccaccggc
 Catggcgctcaaggcgcgcttgcggcgcgacgagaccaacacggtctacaacctggacc
 Gtacctgacggcctcgtggcagttcgccgagcgtggacgctggacgcggggtgcgcta
 Cagcaggtgcgcttcgactcggacgatcattaccaggcgccgggcaacggcgacgacg
 Cggacgcgccacctatcgcaaggccttgcgggtggcgcgctgcgctatgcggccaacga

gaacctgagcctgtacgcctcgtagcgacgaggcttcgagacgcccacgctcaatga

SEQ ID NO:28 polypeptide sequence of Orf14

5 MNTLRRLRILGAAATLGGPAAAEAPAMLEPVRISGTRTGTSLDTPASVDVVDGHELRA
 RNLRQVNLSEGLAGVPGQLQNRQNYAQDLQLSIRGFGARSTFGVRGVRLYVDGIPATMPD
 GQGQTSNIDIGSAGRVEVLRGPFSAALYGNSSGGVVQVFTEQGSDFPEATGSAAAGSFGTW
 RYGAALRGASAADGLDYVLDNFNRTTEGYRDHSAARKNLGNARLGLRMDGSRLLTSLANH
 10 VDLTAQDPLGLTREQFEDDPRSAPVAERFDTRKTVRQTQGGLLYERAFDTRNDLRVMLYY
 GQRRRTTQYQSIIPVAVQQSPTQAGGVIDLGRDYGGADLRWTSRQQVAGLPLTLIGGLAYDT
 MKEQRRGYDNYTGPPAAPTGHGRQGRVAARRDQHGLQPGVPAGLVAVRRALDAGRGAAL
 QHGALRLGRSLPGAGQRRRQRTRHLSQGLAGGGAALCGQREPEPVRLVRLRLRDAHAQ

SEQ ID NO:29 polynucleotide sequence of Orf15

15 Atgaacatgtctctgtcacgcattgtcaaggcggcgcccctgcgcccaccacgctggcc
 Atggcgctggggcgcgctgggcgcgcgccccggcgcatgcccactggaacaaccagtc
 Atcgtcaagaccggtgagcgccagcatccatattcagggtccgacccggcgcg
 gtacggaccgcccagcggaaccaccatcaaggaagcgccgctcaggcccaggcatcctg
 ctagaaaatcccgcggcgagctgcagttccggaacggcagtgctcacgtcgtcgggacag
 20 ttgtccgacgatggcatccggcgctttctgggacccgtcacgtcaaggccggcaagctg
 gtcgccgatcacgccacgctggccaacgttggcgacacctgggacgacgacggcatcgcg
 ctctatgtggccggcggaacaggcccaggccagcatcgccgacagcaccctgcaggcgct
 ggccggcgtgcagatcgagcgcgccgccaatgtcacgggtccaacgcagcgccatcgtcgac
 gggggcttgcatatcgggcgccctgcagtcattgcagccggaagacctccgcccagccgg
 25 gtgggtgctgcgcgacaccaacgtgaccgcccgtgcccggcagcgggcgcccgccggcggtg
 tctgtgttggggggcagtgagcttacgctcgacggcgggcacatcacggcgggcgggcg
 gcggggggtggcgggccatgcaagggcggtcgatcgatcgagcgcgcgacgatacggcg
 ggggacgcgcctgcccggcggtgcggttcccgcggtgcggttcccggtggtgcggttccc
 ggcggcttccggtcccgcggttccggtcccgctcctcgacgggtggtatggcggtggacgta
 30 tcgggctccagcggtggagctcgcccagtcgatcgctcgaggcgccggagctggggcgccga
 atccgggtggggcgggcgccagggtgacgggtgcggggcgccagcttgtccgcaccgcac
 ggcaatgtcatcgagaccggcgccgcgctcgctttgcgcctcaagccgcgcccctgtcg
 atcaccttgcaggccggcgccgatgcccaggggaaagcgctgctgtaccgggtcctgcg
 gagcccggtgaagctgacgctgaccggggcgccgatgcccagggcgacatcgctcgcgacg
 35 gactgcccctccattccggcgacgtcgatcgggcgctcgacgtggcgctggccagccag
 gcccgatggagggcgctaccccgcggtcgactcgctgtccatcgacaacgccacctgg
 gtcacgacggacaactcgaacgtcggtgcgctacgggtggccagcgacggcagcgctcgat
 ttccagcagccggcggaagctggggcggttcaaggtcctgacgggtcaatacgctggcggt
 tcggggctgttccgcatgaatgtcttcgcgacctggggctgagcgacaagctggctc
 40 atgcaggacgccagcgccagcagcagcagcagcagcagcagcagcagcagcagcagcagc
 agcggcaacaccctgctgctggtgcagacgccactaggcagcgcgccgacctttaccctt
 gccaaacaaggacggcaaggtcgatatcggtacctatcgctatcgattggccgccaacggc
 aatgggcagtgagcctggtggcgcggaaggcgccggcgcccaagcccgcgccgcaaccg
 ccgggtcccagccgcccagccgcccagccgcccagccgcccagccgcccagccgcccagccg
 45 ccggcgggcagggagttgtccgcccggcgcccaacggcggggtcaacacgggtgggggtgggc
 ctggccagcacgctctggtacgcccgaagcaatgcgttgtccaagcgccctggcgagttg
 cgccctgaatccggacggcgccggcgccctggggcgcgcccttcgcgcaacggcagcagctg
 gacaaccgcccggggcgccgttcgaccagaaggtggccggcttcgagctggggcgccgac
 cagcggtggcggtggccggcgacgctggcacctggcggggtggccggctatacggcg
 50 ggcgaccgggcttcaccggcgacggcgggcgccacaccgacagcgctgcatgtcgggggc
 tatgccacatatatcgccgacagcggtttctacctggacgagcgtgcgcccagccgc
 ctggagaatgacttcaaggtggcgggcagcgaggggtacgggtcaagggcaagtaccgc
 acccatgggggtggcgccctcgctcgaggcgggcgccgctttaccatgccgacggctgg
 55 ttccctcgagccgagggcgagctggcggtattccggggcgggcggggtgctgctaccgcg
 gccaaaggcctgcgggtgcgacgaaggcgccagctcggtgctgggtcgccggggcctg
 gaggtcgccaaagcgcatcgaactggcaggcgccagggagtgacgcatacatcaaggcc
 agcgtgctgcaggagttcgacggcgcggggtacggtaacacccaacggcatcgcgaccgc
 accgaactgcgcgccgacggcgccgaactgggctgggcatggccgcccgcgctggggcg
 ggccacagcctgtatgcctcgtagctactccaagggcccggaagctggccatgccgtgg

accttccacgcgggctaccggtacagctggttaa

SEQ ID NO:30 polypeptide sequence of Orf15

5 MNMSLSRIVKAAPLRRTTLAMALGALGAAPAAHADWNNQSI VKTGERQHGIHIQGSDDPGG
 VRTASGTTIKVSGRQAQGILLNPAAELQFRNGSVTSSGQLSDDGIRRF LGTVTVKAGKL
 VADHATLANVGD TWDDGIALYVAGEQAQASIADSTLQGAGGVQIERGANVTVQRSAIVD
 10 GGLHIGALQSLQPEDLPSPRVVLRDNTNVTAVPASGAPAAVSVLGASELTLDGGHITGGRA
 AGVAAMQGA VVHLQRATIRRGDAPAGGAVPGGAVPGGAVPGGFGPGGFGPVLDGWYGV DV
 SGSSVELAQSI VEAPELGA AIRVGRGARVTVSGGSL SAPHGNVIETGGARRFAPQAAPLS
 ITLQAGAHAQ GKALLYRVLPEPVKLT LTGGADAQGDIVATELPSIPGTSIGPLDVALASQ
 15 ARWTGATRAVDSLSIDNATWVMTDNSNVGALRLASDGSVDFQQPAEAGR FKVLT VNTLAG
 SGLFRMNVFADLGLSDKL VVMQDASGQHRLWVRNSGSEPASANTLLL VQTPLGSAATFTL
 ANKDGKVDIGTYRYRLAANGNGQWSLVGAKAPPAPKPAPQPGPQPPQPPQPEAPAPQP
 PAGRELSAAANA AVNTGGVGLASTLWYAESNALS KRLGELRLNPDAGGAWGRGFAQRQQL
 DNRAGRRFDQKVAGFELGADHAVAVAGGRWHLGGLAGYTRGDRGFTGDGGGHTDSVHVGG
 YATYIADSGFYLDATLRASRL ENDFKVAGSDGYAVKGKYRTHGV GASLEAGRRFTHADGW
 FLEPQAE LAVFRAGG GAYRAANGLRVRDEGGSSVLGRLGLEVGKRIELAGGRQVQPYIKA
 SVLQEFDGAGTVHTNGIAHRTEL RGTRELGLGMAAALGRGHS LYASYEYSKGP KLAMPW
 TFHAGYRYSW

SEQ ID NO:31 polypeptide sequence of Orf16

Atggcaggacaagcgaggggatgggtacggcgaggcgagcgccacccaatacatatttcaa
 Atttcggcgggcgctgctgctgatgctgggcctgctggacgtcgccggcgccgctgtc
 25 Acggcagcgagcgcaatagatggcgggcgggcggtttctggcgatgtcgccatagcgacg
 Accaaggcgctccgagcagcggtatcaactgactggcgcgacggcgagaggttcgggtgacg
 Gcgggcaccatacgggacgagcggaaccaggcccggttgcgggtcggcacgggagaat
 Gcaccggacaacaccgcgctggcgcgctcggtcttttgcagaacctgatcatcgagact
 Tccgggacggggcattggcgctctctgtccacgagccacagggaggaggaggcagcgcg
 30 ttgtccatgtccgggacgagcggtgcgacgcgcgggcgatgacagtttcgcccgtcgagctt
 tcaggggcctgccagcgccaccttgaaatgacgtggcgctggagacggccggccagcaggcg
 cccgcggtggtgctgtggcaaggcgacagttgaacgcacaggggctggtggttcaggtc
 aacggggcaggcggtttccgcatatcatgcccaggatgccggcagcttcacgttgtcgggc
 tcggatattaccgccccggggcctggaagtgcggggatctatgtgcaggaaggcatgcag
 35 gggacgttgacgggtacgcggtcagcagcgagggcgataccgcgcccgccttgaggtg
 gaggacgcggttacgacgtcagcatgaacggcgggcggtgtgcagacctccggcgcgaaat
 tcgcccgtgcatggctgctggctggcggttcgcgcgagttccgcatagcgtattgagg
 accgtcgagcgagggcctcgcatggcggtggcgctcgctgcgcacagcgaggtcgaactggcg
 catgcgaggtgcgggcgacgggcaaggggctcatggcctggtggtgacgcgaagcagc
 40 gcgatggtgcgggcggttcactggtagagagcaccggagacggcgccgcgcgctgctg
 gaaagcgggcatcttacggtggacggcgagcggtccatggccacggcgcgggcggttg
 gaggtcgacggcgagagtaattgtgtccctgctcaacggcgacgcctgtcgtcgaccag
 ccgacggcgatcaggctgatcgaccctcggtcggtcctgaacctcgacatcaaggaccgg
 45 gcgcagctattggggcgacattgcgccagaggcgagcagcgagcggttcgcccagcag
 gccagggttcgtgtggcgctcgccgacgggggacgtggcgggcgccgacggacggcgcg
 gtccatacgggtgcatgctcgatcgtggcgctctggaccgtgacggcgattcccgggtg
 gccgaggtcaagctggagggcgggcagcgctggcggttgccgacacctgcgagcccaagggc
 gctttcaagacactggtcgcgacgcagggcatttcgggtacgggcacgatagtcatgaat
 50 gcactgttggcgagcgacggcggtgctgctggtggcgccgagggatttcggcgacggg
 caggtgctgggtggtcaacaacacgggatgatggcaccgagagcgggcgacccaagggtg
 ctgatcgaagacgaacaaggccatacggcggttcacgctgggcaacatggggggacgggtg
 gacgcggtgctgctgagctacgaattgaccgagcgagggcgagggcgacgaaggccgc
 acctggcagctgacgcccaggaacgagttgtccaccacggcgacggcgccgctgaatg
 55 atggcgatcgcgcgctcgacgcatctggcaggcggaatggacgtgtgctgcgccat
 atgagcggtcgatcgatcggtcgcgggcggttctggggcgcgggcctgagccag
 cgccagaggtcgataccgggttacggaccctggcagaagcagacgtcagcggaatagag
 ctggggcctcgacaggcggtggcgggcgggcgcaacgacggcggtcggtcggtcgatgctg
 gccggctacagcgagaccggcgcgatggcgggcgcataccgcggcggtgctgacagc

SEQ ID NO:37 polyprotein
MAGQARGWYGAGGRHP IHFQISAGAALMLGLLDVAGAAAVTAAQR IDGGAFLGDVAIAT
TKASEHGINVTRGRTAEVRVTGGTIRTSNGNQAGLRVGTENAPDNTALGASVFLQNL I IET
SGTGALGVSVHPEQGQGGTRLSSMSGTTVTRGDDSFALQLSGPASATLNDVALETAGQQA
PAVVWLQGAQLNAQGLVVQVNGAGVSAIHAQDAGSF T LSGSDITARGLEVAGIYVQEGMQ
GTLTGTRVTTQGD TAPALQVEDAGTHVSMNGGALSTSGANS PAAWLLAGGSAQFRD T VLR
TVGEASHGVDVAAHSEVELAHAQVRADGQGAHGLV VTRSSAMVRAGSLVESTGDA AALL
ESGHLTVDGSVVHGHGAAGLEV DGESNVSLNGARLS SDQPTAIR LIDPRSVLND I KDR
AQLLGD I APEAQQPDGSP EQARVRLADGGT WAGRTD GAVHTVRL LDRGVWTVTGDSRV
AEVKLEGGTLA FAPPAQPKGAFKTLVATQGISGTGTIVMNAHLP SGTADVLVAPQGFGR
QVLVVNNNTDDGTESGATKVPL I EDEQGHTAFTLGNMGG RVDAGARQYELTASEAQADKAR
TWQLTPTNELSTTTATAAVNAMAIAASQRIWQAEMDVLLRHMSGLHS I GSPGGFWARGLSQ
RQR LDTGYGPWQKQTVSGIELGLDRRVAGGATTAWSVGMLAGYSETRRDGGAYRAGHVSH
AHVGAYVSYLNDSGSYVDGVVKYNRFRHGFDIRTTDLKRVD AKHRSHGLGALLRGRR.ID
IDGGWYVEPQASVAFWHAGGSRYEASNLGRVADGASHVWLRAHGA EAGRQMR L ANGNIVE
PYARLGWAQELGADNAVYNTNGIRHVTRSRGGFAEARVGVGALLGKRHALYADYEYAKGAR
FEAPWTLGLGYRYSW

Atgtatctcgcgatagattccgctcaatgtccgctcttcccttgcagatccgcggttccgcgctgg
 Cgcctgcatacgcctggccgcgcagctctggcgctggccggcatggcccggtggcgcccgcg
 Gcggcgcaggcgccgcgcagcgcctgtggccggcgcgccgcgatgcgccaggacgcccgggcag
 Gaaggagagttcgaccaccgggacaacacgctcattgcagttctttgacgcagggcgtcggc
 Atcaatctcgacgcagcatcccgcgcagctcggcgagacggcgccccccacgctcaaggac
 Atccacatctcgggtggagcacaagaaccgcgatgagcaagccggccatcgggtgctgctgc
 Agcgggcgccggccgcgcgctgacgctggccggctcgaccatcgatgccaccgagggcggc
 Attcccgagtggtacggcgcgccgcgcagctggagatctggtatggcgtcacctggcggggc
 Ggggaaggagtgagagccgatcaggtctctgacgcggcgagcgcctgagcgtgcgcggc
 Ggcgtgctggcgggcggaagcgccggggcgtcggcctggtccggggccgcgcaaggcgggccag
 Gcgagcatcatcgacgcgacgctgcagagcatcctcggggccgcgctcattgccgacggc
 Ggctccatttccgtcgccggcggttcgatcgacatggacatgggcccgggattcccgcgg
 Ccgctccaccgcttcccggggcgccgcgctggccgcgcgatccgcgctcgatcgcgttgc
 Gcggtgcacgcggccaggacggcgaaggtgacactcggggaggtggcgctgcgggctcac
 Gggccgcaggcgacggggcgtctatgcgtatatgcctggcagcgaatcacctgcaggga
 Ggcacagtcagcgtgcaggggcgatgacggggccggcgctggtcgccggcgcgggcctgctc
 Gacgccttgcgcgccggggcgacgggtgcggctggacggaaccacgggtgtcgaccgatggc
 Gccaacaccgatgccgtgctgggttcgcgggcgacgcgggcgcgcgcagggtcgtcaacac
 Gtgctgcgcaccgcgaagagcctggccgcggcgatcggcccgagcatggaggcccgcgctc
 Acgctgcgggcagaccgcgatcgagacgcggcgcgggggccgagggcatctccgtgctg
 Ggcttcgagcgcagctccggctcgcggccggccagcgtcgacatgcaggggcggcagcatc
 Accacgacccggcaaccgcgcggccggcatcgcgctcacccacggcagcgcccgcctggaa
 Ggcgtggcggtgcgcgccgagggcagcggtcgagcgccgcgcagctggccaacgggcacg
 Ctggctgctcagcgcagggtcgctggcctcggcccgatcggcgcgatcagcgtgaccgac
 Acgccgctgaagctgatcgggggcgcccctggcgacgacagcggctcggctcgggttgacc
 Gacggcgccacggcgcaaggcggaatggcggttctctccagcagcatccaccattccg
 Gtggcggttgccctcgagagcggcgccctggctcgcggcgatatcgtcgccgacggcaac
 Aagccctcgatgccgggatctccctcagcgtggccagcggcgccgcctggcacggcgcc

Acccaggtgctccagtcggccacgctgggcaagggcggaacctgggtcgtgaacgcgcgac
 Tccccgggtgcaggacatgtcgatgcgcggcgggcggtcgaattccaggcgcccgcgccc
 Gaggcctcttacaagacctgacctgcaaaccctggacggcaacggcggtgttcgtgctg
 Aacaccaacgtcgccgcccgggcagaaacgaccagttgcgggtcacggcgccgcatggc
 5 Cagcaccgcgtgctggtgcgcaatgccggaggcgaggccgacagccggggcgcccgctg
 Ggcctggtgcatacccaggggcagggcaacgccaccttccggctggccaacgtcggcaag
 Gcggttgacctgggacgtggcgctacagcctggcgaggatccgaagacgcatgtctgg
 Agcttgacgcgcgggccaggccctgtcgggggcgggccaatgccgccgtgaacgcggcg
 Gatctttccagcatcgccctggccgagtcgaacgcgtggacaagcgctgggagagctg
 10 Cgctgcccgcgcgacgccggcgccatgggcgcgtacgttcagcgagcgccagcagatc
 Agcaaccgcccacgcccgcgcctacgaccagacgggtcagcgggctggagatcgccctggac
 Cgtggctggagcgcgctcgggcgggcgctggtacgcggcgccctgctcggtacacctat
 Gccgaccgcacctatcccggcgacgggtggcggaaggtcaagggcctgcacgtcggcggc
 Tacgcccgcctatgtcggcgatggcggtactatctcgacaccgtgctcggtcgggccgc
 15 Tacgatcagcaatacaacattgccggcaccgatggcgggcgcgctcaccgcccactaccgc
 Acaagcgccgcgcgcatggctcgctcgaaggcgggcgccggttcgagctgcccacgactgg
 Ttcgcccgaaccgcaggccgaggtcatgctgtggcgacgtcaggcaagcgctatcgcgcc
 Agcaatggcctgcgcgtcaaggtggacgccaacaccgccacgtgggcccctgggcttg
 Cgcttcggccgcccgcacgtcgccctggccggcggaacatcggtgcagccctacgccaggctc
 20 Ggctggacgcaggagttcaaaagcacggcgatgtgcgccaccaatggcattggccatgcc
 Ggcgcaggccgcccacggccgcgtggaactggcgcgggcgctcgacgcgcgcttgggcaag
 Gggcacaacctctatgcttcgtacgagtagcgggcgaccggatcaacattccgtgg
 tcgttccacgcccggctaccgctacagcttctga

25 SEQ ID NO:34 polypeptide sequence of Orf17

MYLDRFRQCPSSLQIPRSARLHALAALALAGMARLAPAAAQAPQPPVAGAPHAQDAGO
 EGEFDHRDNTLIAVFDGGINLDDDPDELGETAPPTLKDIIHISVEHKNPMSKPAIGVRV
 SGAGRALTLAGSTIDATEGGIPAVVRRGGTLELDGVTVAGGEGMEPMTVSDAGSRLSVRG
 30 GVLGGEAPGVGLVRAAQGGQASIIDATLQSLILPALIADGGSISVAGGSIDMDMGPGFPP
 PPPPLPGAPLAHPPPLDRVAHVHAGQDGKVTLEVALRAHGPQATGVYAYMPGSEITLQG
 GTVSVQDDGAGVVAGAGLLDALPPGGTVRLDGTTVSTDGANTDAVLVRGDAARAENVNT
 VLRTAKSLAAGVSAQHGGRTVLRQTRIIETAGAGAEGISVLGFEPQSGSGPASVDMQGGSI
 TTTGNRAAGIALTHGSARLEGVAVRAEGSGSSAAQLANGTLVVSAGSLASAQSGAISVTD
 35 TPLKLMFGALASSTVSVRLTDGATAQGGNGVFLQQHSTIPVAVALESGALARGDIVADGN
 KPLDAGISLSVASGAWHGATQVLQSATLGKGGTWVNVNADSRVQDMSMRGGRVEFQAPAP
 EASYKTLTLQTLTDGNGVFLNTNVAAGQNDQLRVTGRADGQHRVLRNAGGEADSRGARL
 GLVHTQGQGNATFRLANVGKAVDLGTWRYSLAEDPKTHVWSLQRAQALSGAANAAVNAA
 DLSSIALAESNALDKRLGELRLRADAGGPWARTFSERQQISNRHARAYDQTVSGLEIGLD
 40 RGWSASGGRWYAGLLGYTYADRTYPGDGGGKVKGLHVGGYAAVYVDGGYYLDTVLRLGR
 YDQQYNIAGTDGGRVTADYRTSGAAWSLEGGRRFELPNDWFAEPQAEVMLWRTSGKRYRA
 SNGLRVKVDANTATLGLRLFRFRRIALAGGNIVQPYARLWGTQEFKSTGDVRTNGIGHA
 GAGRHRGVELGAGVDAALGKGHNLASYEYAGDRINIPWSFHAGYRYSF

SEQ ID NO:35 polynucleotide sequence of Orf18

45 Atgcacatttacggaaatatgaatcgagcaacaccttgccgcggtgcggtgcgtgcctt
 Gcgcttgccctgctgggagcgggtatgtggacactttctcctccctcggcattggcgctt
 Aagctcccgtcgctgctgacggacgacgagctgaagctggttctgcccactggcatgtct
 ctggaggatttcaagcgcagccttcaggagtcgcgcgcgagcgcgctggcaacgcgcgcg
 50 tcgtcttcgcctccggttgcaagccaggtccgggtccgttgccgaggtccgtcgggg
 tcggggccacaaggacaacccatcccctccgctcgctcgccgctcggtccaggtatggcggg
 tcgtctggcgacataacccccggcggtgggggggggacgcataaaaaatgggttgcccgg
 ataggaaaggtcggcggggtctgcgcccggacggataccagtaggggtcgggtcccgcac
 gccggcatggctccggagcgggttcgagctcgcccgccgcatcggtggggcgggcaag
 gatgcgatgccgccctcggaaggcgagaggccggactccggtatgtccgattcggggcg
 55 ggtggcgaatcgctcggtggaggttgaaatccggacggcgctggcaagccaccgcgggag
 gaagggcagccgggttccaagtctcctgcggacgggtggccaggacggggccgcccggccc
 cgggacggcgcgatgcggatccgcaacctccgcgggacgatggcaatggggaacagcag
 ccgcccgaagggcgcggggatgaggggcagcgcccgcgcctgcgcgcggaaacgggtggc

aacggtggcaacggaacgcgcagcttcccgagcgcgccgacgacgcggtccgaagcct
 cccgagggagagggcgccgatgaaggtccgcaaccgcccagggcgccgagcaggac
 ggcgcggaggttccctcccgctcgccgcggcgccgcccgggcaacggtgtctatgacccg
 5 ggcacgcataccttgaccacgcgcgctctgcgccggtgagcctggccagcagttcgcat
 ggcgtatggcaggccgagatgaacgcgttgagcaagcgcagtgccgagttgcccctgacg
 ccggttgccggcgccgctatggggcgcgcttttgcccgccgagcagtcgacaaccgc
 gtgtcgcgaggttccgcccagaccatcagcggtttcgaactggcgccgataaccgcttg
 ccggtggccgacggcgctggcacgtggcgcggtggctggctacaccaacggccgcatc
 10 aagttcgaccggggcgccacggcgatgacgacagcgtgcacgtggcgcttacgctacc
 tacatcgaggacggcggtttctatatggatggcatcgtgcggtcagccgcattcgccac
 gcgttcaaggtggacgacgccaaggccggcgctgcccggccagtagccgccaatggc
 gtggcgcgctcgctggaactgggcaagcgcttcacgtggcccgccgctggtacgtggag
 ccgagctggaggtggccgcttccatgcgcaaggggcccagctacaccgcccagcaacggc
 ctgcccacaaaggacgacggcacgaactccatgctggcgccgctggcgctgacgtggg
 15 cggcagttcgacctggcgatggccgctggtgacgcccctacatgaagctgagctgggtg
 caggagttcgacggcgaaggcgacgggtgcgcaccaacgacatccggcacaaggtgcggctc
 gatggcgcccgacccaactggcgtaggggtggcttcgcaactgggcaagcacggcagc
 ctgttcggctcgtagcagtagcgaaggcgagccgcccagaccatgcccgtggacctccac
 20 gtcggctatcgctacgcctggtag

SEQ ID NO:36 polypeptide sequence of Orf18

MHIYGNMNRATPCRGAVRALALALLGAGMWTLSPPSAWALKLPSLLTDELKLVLP TGMS
 LEDFKRSLQESAPALATPPSSSPVAKPGPGSVAEAPSGSGHKDNPSPPVGVGPGMAE
 25 SSGGHNPGVGGGTHENGLPGIGKVGSAPGPDSTSGSPDAGMASGAGSTSPGASGGAGK
 DAMPPSEGERPDSGMSDSGRGGESSAGGLNPDGAGKPPREEGEPGSKSPADGGQDGP PPP
 RDGGDADFPPRDDNGEQPPKGGGDEGQRPPPAAGNGGNGGNGNAQLPERGDDAGPKP
 PEGEGGDEGPQPPQGGGEQDAPEVPPVAPAPPAGNGVYDPGTHLTTPASAAVSLASSSH
 GVWQAEMNALSKRMGELRLTPVAGVWGRAFRQRQVDNRSREFRQTISGFELGADTAL
 30 PVADGRWHVGAVAGYTNGRIKFDRGGTGDDDSVHVGAATYIEDGGFYMDGIVRVSRIRH
 AFKVDDAKGRRVRGQYRNGV GASLELGKRFTWPGAWYVEPQLEVAAFHAQGADYTASNG
 LRIKDDGTNSMLGRLGLHVGRQFDLGDGRVVQPYMKLSWVQEFDGKGTVRTNDIRHKVRL
 DGRTELA VGVASQLGKHGSLFGSYEYAKGSRQTMPWTFHVGYRYAW

SEQ ID NO:37 polynucleotide sequence of Orf19

35 Atgaaaccgacttccatcctggcacgtttgccccgctatctcgccgctgcccgcgtggcc
 Gcgtggccgcccgtggtgtcgccgctcgccgcccgcagggcacagactccgctgccc
 Gcgggactcggcgccgcccaggtgcccagtttgtccggcctgccgtccgatgccctg
 cgccagcaggcgtcggtggctggcgccgcccgtgttgccgcccctatctgtcaggcctgacg
 40 gatgcgcaattgcccgaatatgtgcaggcgtgacacccgggcagatcacgcaggggctg
 gcggcgttgacgcctgcccagcgtgcccgggtgcagcgcgaattcgaacggcaggcgccg
 cggcaggtgcagcaggcggtagggccgaggtcgcccgccgagcgcgcggcggtggcg
 atggggcagagcgcacatcgatgctgctcgacgccgagatgggaaccctggcgcaacgc
 cagggcgatctgcccgcggccacgacgagggcgcccttctggcgccgagcgcgaac
 45 cgcttcaaggtcgatagccggacacaccggcgttcgacctgcccgtggagtacctgacg
 ctggcgcccaccatggctggcgccctggacacggggcggtctatctggcgccctacgcc
 ggcgtctcgccgcccgcagtgatgacaacgacatcatgcacggccggatcgaaagccgg
 ttcttgggcacgtacctgacttatgtggacaacggcggttctacgtcgatgcccgtcagc
 aagctggggcgctacgacgagtcctgtgcttcgacctgccgctggggctggcgactac
 50 gacgacgatataatcgatacaacgtatacgggcagtgccgagggccggctatcacttcaag
 ttgccgcaacgctggttcgtcgagccgcaggcgaggtgatctactcgccgagcagccag
 acgtcggtgcagggggcgccggcggtgcccgcggcgccgggatttcacctggccggcgcc
 ggcacctgctgcttctatgtcagcgcctcgtagctgacgagttctcgacgacgactcg
 gtcgatttcggcggaagtcgtacgatgcccgaactgcccggcagccgctggcagctgggt
 55 gccggcgccggcgctggacgtggggggcgcatcgccgctacgcggatctgcgctatgggcac
 ggcgccaacatcagccaggacctgtcgctgaacatcggctacgcgtaccgcttctag

SEQ ID NO:38 polypeptide sequence of Orf19

MKPTSILARLPRYLGCALAALAAVAPLAPAQAQTPLPAGLGAAEVRQYLSGLPSDAL
 RQQASWLAPALLRPYLSGLTDAQLRQYVQALTPGQITQGLAALTPAQRARLQREFERQAR
 RQVQQAQVRAEVAARSARAVAMGQSASMLLLDAEMGTLAQRQGDLLRRGHDEGAFWARGSAN
 RFKVDTPDTPAFDLRVEYLTGADHWRLDTGRLYLGAAYGVSRRMDDNDIMHGRIESR
 5 FLGTYLTYVDNNGGFYVDAVSKLGRIDESVSFDLPLGLGDYDDDI SHTTYTGSAEAGYHFK
 LPQRWFVEPQAQVIYSRSSQTSVQGRAGVRAGRDFTLAGGATLRPYVSASYLHEFSHDDS
 VDFGGKSYDAELPGSRWQLGAGAALDVGAHRAYADLRYGHGANISQDLSLNIGYAYRF

SEQ ID NO:39 polynucleotide sequence of Orf20

10 AtgggtcggcaggagttgtcatcgtgcaggggtggttataccggggcgaccttctctTtatac
 GccgcaaactgtgcttatcccgcggaacgccaatcggtttctgggtccgggtcagGtgtcg
 Aacgggcccagtcacgtcaccgcactgggtgggtgggcccgggaactgatcgtcggggatcacg
 Ggcggccgggacctgtctcatcgaggccggcggtaccgtgctcaacgactgggacctatctc
 Ggcagtgacaatggcgctgtgggcaccctgacgggtgtcgggcccgcgacggcgccggggcc
 15 Gcgtcgacctggagcactgtcgacgatgtgtcgatcggcggttgccggcgggcagcaggggc
 Acgctcgaggtgctcggcgggggccagggcgcaaagcggtatggggcaccatcggcgtcgct
 Gcaggcagcgtcggaagcgtgacgggtgtcggggccggggtcggtgtggaatatcgccacg
 Gtcaattcgttccagatcggtcgggcccgcagcgggacgctgtggatcgaccagggcggc
 Gcagtgtagcagggcagggcgctcatcggttggaaacccggcagcgacgggcacgtcacg
 20 Gtattgggtccggcaacgggtatggaaccgctgaacaatatctatgtcggtctcggcggg
 Actggtgaactggatatccgggacggcgcgccggttgccgactgcaggggtcgagcccgcg
 Ggcggccggcgatcgatctacatcgggacgagcgagggagcgccgggcacggtaacgggtg
 Tcgagcgcgacggcgctcacctcgacgctcacgtcgaccgaccgtatcgaaatcggtcgtc
 Gcggcgccgggggtgctgactgtcgccaaaggcggtatgggtggcgctcgccagcgacgccc
 25 Tggtagccatcacgggcacgtcctcgggaacgctgaacctgaccggcgatgccagcggc
 Cgaggcggtgctggaaacggggtcggtcatcaaggcgccgggcaacgcgaccttcaacctg
 Gatggcggggtcctgcgcgccaatcggtgacgaggccaatttccctcaatgggtttctccacg
 Caggcggtgggaagcgggcgccgctgggttcgatagcaatgcccatgacgtgggctgtgtc
 Accgcttctcgggtacgtccagcttcaacaagctgggagccgggcacgctgacgctgtca
 30 Ggcaacagcgccgcttcacggggaaacaccgatccaggccggaacgctgacggtggac
 Ggcgttcttggcgggccgggtggatgtgctggcgggggcgcggttgaccgggtaccggggcg
 Gtcggtgcgacggccaacaagggcaccattgcgccggggcccgcgacgggcttggcacc
 Ctgacgatcgccggggattacgcgggccaggcgggcaacctggaaatccgtacgcagctt
 Ggcgccgacgactcgccgaccgacagggtgggtgatcacgggcccagcgctggcggtgaca
 35 Cgggtcacggtcgagaatatcggcggcagggcgccctcgacccagcggggcatacaggtc
 Gtcgaggtcaatggcgcttcggcgaggcggttcaacctcgccaacggcgattacgtcatc
 Gaggggcgctccggcgctgggtggcgggcgccctatgggtatgtgctgcagcaggacggcgcc
 Gacggcgattgggtatctgaaatcgctcgctgcccgaacctgggggtccccaaggcgggggc
 Ggtctgcccggcgccggggagcccgtgctttatcagcccggcggtgccgggtctatgaagcc
 40 Tatgccaaacagctgctgcatctgagccgggtttccaccttgcgccaaacgggtcggcaat
 Cgcctttatgatccggcagatgtcggcggcaacggcggtatggagccgctcgagggtcc
 Gcgagccagctcgatccttccggtccacgactggcgaaacggcaggacgtcgatagctgg
 Aaagtgcagttcggtgtcgaccgtatcctggccggcggggaagagggtcccgcctgggtg
 Ggcggattggcgctgcagtagcggaaggccgacacgcgctgtcgatatacggcaat
 45 Ggcactgtcgacgccacggcctatggcctgaccccgacgctgacctggtacggcagggac
 Ggcgcctatgtcgatgcccaggcccaggcgatctggttcgacagcgacctgagttcacgg
 Ctgcccggaagctcaaggatggcggaaggcgcatgggtatgggctgggtatcgaagcg
 Ggcaaggccttcggattgcccgggggggtggccctgatcccgaggcgcaattgtcgtac
 Gcatcgaccgcttcgacagcttcgacgacagattcgccgcccgcgtcgaagacgataag
 50 Ggcgacagcctgcagggcgctctcgccatcgcgctggactacaagagcagctggcaagcg
 Ggcggcgcgaaacggggagtcgagtgctcttcggcatcgctcaatgtgaagcatgagttcctg
 Gatggcacgcgctgctgctgttcggcggtgcccggtaagcagccgcatggcgcgcacctgg
 Ggcagctgggagtgggggcggtatcggttggggagaacgctacgcatttacggccag
 55 Gtgacgcccgatgcagatttcgcccgcagctacatcgctcaccgcgaccgcgggggttcagg
 atgatgttctag

SEQ ID NO:40 polypeptide sequence of Orf20

MVGRSCHRAGWLYRATFLLYAANCAYPANAQSVSGSQVSNGPITSPHWVVGELIVGDT

5 GAGTLLIEAGGTVLNDWAYIGSDNGAVGTLTVSGRDGAGAASTWTTVDDVSIGVAAGSRG
 TLEVLGGARAQSGWGTIGVAAGSVGSVTVSGPGSVWNIATVNSFQIGSGSGTLWIDQGG
 AVYSGQGVIGWNPNGSDGHVTVLGPATVWNPLNNIYVGLGGTGELDIRDGA AVATAGSSPP
 10 GAAASIYIGTSAGSAGTVTVSSATAVTSTLTSTDRIEIGSAGAGVLTVAKGGMVGVASDA
 WIAITGTSSGTLNLTGDASGRGVLETGSVIKAGNATFNLDGGVLRANRDEANFLNGFST
 QAVGSGGAWFDTNAHDVGVVTAFTSGTSSFNKLGAGTTLTSGNSAAFTGNTDIOAGTLQVD
 GVLGGPVDVLAGARLTGTGRVGATANKGTIAPGPRSGFGTLTIAGDYAAQGGNLEIRTQL
 GADDSPTDRLVITGASAGVTPVTVENIGGTGASTQRGIQVVQVNGASAGRNFNLANGDYVI
 15 EGRPALVAGAYGYVLQQDAADGDWYLKSSLPDPGAPQGGGGLPGAGEPVLYQPGVPVYEA
 YANTLLHLSRLSTLRQRVGNRLYDPADVGRNGVWSRVEGSASQLDPSASTTGERQDVDSW
 KVQFGVDRILAGGQEGSRLVGGGLALQYGKADTRVSSIYNGTVDATAYGLTPTLTWYGRD
 GAYVDAQAQAIWFSDLSRLAGKLKDKRAHGYGLGIEAGKAFGLREGLALIPQAQLSY
 ASTRFDSFDDRFGARVEDDKGDSLQGRGLGIALDYKSSWQAGGANRESSVFGIVNVKHEFL
 DGTRVRVAGVPVSSRMARTWGSVGVGADYGWGERYAIYQVDADADFAGSYIVTATAGFR
 MMF

SEQ ID NO:41 polynucleotide sequence of Orf21

20 Atgccgtcacccgatgccttgccgcacacgccgcctgcttcaggcggcgatcgcgatgatc
 Agcgggatcctgcagcaggacctcgccagttggctggcgccggatgccgcaaagcgagc
 Ccctccgagcctggcaaggcgccgaaaaaatcggggtaatgccgaacgaggacctcgcc
 Aagtggctggttcggggggcgaaaaagaacaatccgcccagcctggcaagacgctggac
 GaaatccgTgcgggtctcgaaaaatgggtggcgcccggtccaagccgcccgtcgaaaccg
 Gatccggacaaggcgacgcaggcgtatcgcaaaagacctcgataaatggctggcgccctccg
 25 Gccaagtccggcccgccgaagcgccacccgtcgtccaacccgaagcgcccgccgaagcg
 Caacctgaggcgccgctgtcgtgccgcccgcggccgagccgcccagcagctcgaccgccg
 Gccgttcgcccgcgcggccggcgccggcgacgcgggtgtacgtgccgggacgcgcacgctg
 Acgccgacggccaacgcggcggtgggcacggccagcgccgcgcaaggtctgtggcaggcc
 Gagatgaacgcgttgagcaagcgcatgggcgagttgcgccctgacgcgggttgcgggcggc
 30 Gtatggggcccgcttttggccggcgccaggacgtcgacaaccgcgtgtcgcgcgagttc
 Cgccagaccatcagcgggtttcgaaactgggcgcccataccgccttgccgggtggccgacggg
 Cgctggcacgtgggcgcggtggctggctacaccaacggccgcacatcaagttcgaccggggc
 Ggcacgggcatgacgacagcgtgcacgtgggcgcttacgctacctacatcgaggacggc
 Ggtttctatatggatggcatcgtgcccgtcagccgcattcgccacgcgttcaaggtggac
 Gacgccaaggccggcgcgctgcccggcagtcaccgcccgaatggcggtgggcgcgctcgctg
 35 Gaactgggcaagcgcttcacgtggcccgccgctggtagcgtggagccgcagctggaggtg
 Gccgccttccatgcgcaaggggcccactacaccgcccagcaacggcctgcgcacatcaaggac
 Gacggcacgaactccatgctgggcccgcctgggcctgcacgtggggcgccagttcgacctg
 Ggcgatggccgctggtgcagccctacatgaagctgagctgggtgcaggagttcgaccggc
 Aagggcacggtgcgcaccaacgacatccggcacaaggtgcggctcgatggcgcccgacc
 40 Gaactggccgtaggggtggcttcgcaactgggcaagcacggcagcctgttcggctcgta
 Gagtacgccaagggcagccgcagaccatgccgtggaccttcacgctcggtatcgctac
 gcctggtag

SEQ ID NO:42 polypeptide sequence of Orf21

45 MPSPDALPHTPPASGGDRVISGILQDLGSLAPDAAKRSPSEPGKAAEKIGVMPNEDLG
 KWLVPGAQKNPPEPGKTLDEIRAGLEKQWVAPGSKPPVEPDPKATQAYRKDLKWLAPP
 AKSGPPEAPPVVQPEAPPQAQPEAPPVPPPAEPPAARPPAVPPARFAGDAVYVPGTRTL
 TPTANAAVGTASAAQGLWQAEMNALSKRMGELRLTPVAGGVWGRAFGRRQDVNDRVSREF
 50 RQTISGFELGADTALPVADGRWHVGA VAGYTNGRIKFDRGGTGDDDSVHVGA YATYIEDG
 GFYMDGIVRVSRIRHAFKVVDAKGRVRGQYRGNGVGASLELGKRFTWPGAWYVEPQLEV
 AAFHAQAGADYTASNGLRIKDDGTNSMLGRLGLHVGRQFDLGDGRVVQPYMKLSWVQEFDG
 KGTVRTNDIRHKVRLDGGRTTELAVGVASQLGKHGSLFGSYEYAKGSRQTMPWTFHVG YRY
 AW

55 SEQ ID NO:43 polynucleotide sequence of Orf22

Atgtgcgacacctgcagagatgatgatggcacctcgcttcgattcgcgctccaaggcggg
 Gttgttcaggggcgccatgggtgcaataacgtcgctgtggtggcaacagggtctggaaag

Gtcgcatcgagaatgcggaactgctcggagccagcggcatgtacgccacgttcggcgcg
 caggtcgatatgaaagggggcgcatctctggcgcacacaaccaatatcctgggaagccag
 gggtacgccgatgggtccctatggcgggcggtcggtgacagaggacgggtcaagtcaacctg
 gagggcgccaaggtcagtgcaactggcctggggggccgcccggcttggtggtgctggggcgac
 5 aaggacaccagcccgcgagccagcctgcgcaacaccgacgtccacggagaggtcgccgcc
 attgcgctgggggttcaatggcgaggcgcaacatctcggcgggcagcttgagcgtagaggat
 gggggcgtgctcaccaccctgacgcccgatgcagtcgagttactactacgactacgccttg
 tccatggagcatctgccagctgatgcgccgttgacgcccgtccgcgtcacgctgtccgat
 10 ggcgcgcgcgccagcggagaaacgttgatcgcgcatggcgggttggtgcccacgacgctg
 cgcttgagcagcggggtcgacgcccgcggcgacatcgtcacgctgccgccttccgcgcgcg
 ccgattccgcggagcaaccggatgccgagccggaaccggatgccgagctggaaccggac
 gccgcggcgagtcggacgccaaggcggaatgcgcgggtcatggcgaggtagatggcggg
 gaacctgttgccgtgccgatcccggccccttcgcatcccgatgcccgatcgacgtgttc
 atcgacagcgggtgcccaatggcgggggcatgaccaagaccgtcaatgcgttgcgcatcgag
 15 gacggcacctggaccgtcaccgggtcgctccacgggtgaacagcctgcacctgcaggcaggc
 aaggtggcgtagcgaacgcctgccgaaagcgagcggagaattcaaacacctgcgggtcaag
 accctctcgggaagcggcctgttcgagatgaacgccagcgccgacctgagcgtggcgac
 ctgctgggtcggtccgacgagggccagcgggcagcacaaggtgctggtgcgaggagccggc
 acggaaccaccgggtgtggaaagcctgacgctgggtcgagctgcccgagggcgagccagacg
 20 aagttcacgcttgccaaccggggcggggtgggtcgacgcccgcgcgttccgctatcgccctg
 acgcccggacaacgggtgtctggggcctggaacggaccagccagcttccggcgtcgccaac
 gcggccttgaataccggggcggtggcgccgcccagcagcatctgggtatgcggaaggcaat
 gcgctctccaagcgctggggcgagttgcggctcgatcccggcgggcgggcgttctggggg
 25 cgcacgttcgcccagaagcagcagctcgacaacaaggctggccgacgcttcgaccagaag
 gtgtacgggttcgagctggggggccgaccatgccatcgaggacagcaagggcgctggcac
 gtggggcgccctgctgggctataccgcgcgaaggcgagcttcatcgatgacggcgccggg
 cataccgacagcgcgcatatcggggcctacgcggcgtagctggcggaacaacggcttctat
 ttcgattcgaccctgcgcgcgcccagccgttcgagaacgacttcacggtaacggccaccgac
 30 gccgtttccgtacggggcaagtaccggggccaatggggtagggcgccaccttgaggggcggc
 aaacgtttcacgctgcacgacggctgggttcgtcgaaacctcagtcagggtgtcgctgttc
 catgccagcggcggaacctaccgtgcgcgcaacaacctgtcggtcaaggacgaaggcggc
 acctccgcgctgctgcgcctgggcttggcgggcgggcgacgcatcgacctgggcaaggac
 cgcgtgatccagccctatgccaccctgagctgggtgcaggaaattcaaaggcgtcacgacc
 35 gttcgccaccaacgggtacgggtgcgcgaccgacctgagcgggtggccgggctgaattggcg
 ctgggcctggccgcccgttggggcgcgggccaccagctctacacttcgtacgagtagcgc
 aagggaacaagctgacctgccttgagcgttccacctgggctatcgctacacctggtag

SEQ ID NO:44 polypeptide sequence of Orf22

MCDTCRDDDGTSPIRVQGGVVQGGMGANNVAVVATGSGKVAIENAEELLGASGMATFGA
 40 QVDMKGGRILAHNTNILGSQGYADGPYGGVVVTEGQVNLEGAKVSATGLGAAGLWLLGD
 KDTSPRASLRNTDVHGEVAIALGFNGEANISGGSLSVEDGAVLTTLTPDAVEYYDYAL
 SMEHLPADAPLTPVRVTLSDGARASGETLIAHGGLLPMTLRLSSGVDARGDIVTLPPSAP
 PDSAEQPD AEPEPD AELEPD AAAQSDAKANARVMAQVDGGEPPVAVPI PAPSHPDAPIDVF
 45 IDSGAQWRGMTKTVNALRIEDGTWTVTGSSTVNSLHLQAGKVAYATPAESDGEFKHLRVK
 TLSGSGLFEMNASADLSDGDLVVSDASGQHKVLRGAGTEPTGVESLTLVELPEGSQT
 KFTLANRGGVVDAGAFRYRLTPDNGVWGLERTSQLSAVANAAALNTGGVGAASSIWYAEGL
 ALSKRLGELRLDPGAGGFWRGTFQAQKQLDNKAGRRFDQKVYGFELGADHAIAGQQGRWH
 VGGLLGYTRARRSFIDDGAGHTDSAHI GAYAA YVADNGFYFDSTLRASRFENDFTVTATD
 50 AVSVRGKYRANGVGATLEAGKRFTLHDGWFVEPQSEVSLFHASGGTYRAANNLSVKDEGG
 TSAVLRGLGLAAGRRLDLGKDRVIQPYATLSWLQEFKGVTTVTRTNGYGLRDLSSGRAELA
 LGLAAALGRGHQLYTSY EYAKGNKLTLPWTFHLGYRYTW

SEQ ID NO:45 polynucleotide sequence of Orf23

Ttgcgccagaccacgcgggtgccggtgcggctcgctcctgcgcggcgcgggcggtcgcgag
 55 Ggcgatgtcgtgcgcgcgcccagacggcgccggagaaggatgggttcggcacgcccgtg
 Cggccgggcttgcgcgtcgggctggaccaggcgccgctcgagctcgatgtggccgacggc
 gcgcagtgccatggcgcgactcagtcgcttgacaggctggccctgggcgcggggcgccaa
 tggcgcatgagcgcggcatccagcgtgggcgaactgagcatggagcctggcgcgggccgtc

5 gtgttcggcgatgcggccggaccgggttttcaaacgctgacggtgcgcaccctggcgggc
 gccggttcgttcgagatgcgtgcggacgccgcgctggagcatgccgatcaactggtggtg
 accgaccaggccgaagggcgccatcgcggtggttgcgcgccgcccggcgccgagccg
 tcaaggcacaggccgtgctggtgcgcgcgcccgcagacggcaaggccagtctcgaactc
 gacggcagcgacggcaggccgacttcggcacctatcgctacgggctggcgcagcagccg
 10 ggcggcgccctggggcctagtacggacggggtattcgctccaccgcccgcggcgctggat
 accggcggaactgggcgcggtgcaggggttctggtatgccgaatccaacgcgttgggcaag
 cgcattgggcgaattgcgcctgaacccggacgccggcgccgctggggccgggcgttcagc
 cagcgccagcgcatcagtcgcgcgcggcgccgcatctccagcaaggcgtcagcgccatc
 gagctgggcgcgcgaccgggcctggccgtggccggcgccgcttggcatgcccggctggttg
 ctgggctacacgcgcgcgtcacgcggttttccggccagggaaggggcacaccgacagc
 gtgcacgtgggcggctatgccacctatcgcgcccaatggcggtgtacgccgatgccacg
 15 ctgcgcgccagccgcttcgagaattcgcttcgacgcacctggctgggcggggcgccaccgtg
 tccggcagctaccgcgcgaatggcggtggcggtgacgctggaggccggccggcgcttgggc
 ctggaccggcactggttcgtcgagccgcaggccgaactggcggtggttctgctgcggcgcc
 ggtacgtacacggccagcaatggcctgcgtatcgaggatgacggcgccacgtcgctgcag
 gcgcgggttaggcgcgcaagccggccgcgcttcgacttgccgcggcgccggtggtgcag
 ccctacgcgcagctgagttgggtgcaggaactcaaggcggtgagcacggtgcgcaccaac
 20 ggcacgcgcacgtaccgacctgggcgcggggccgcgtcgaactgggactgggcgtggcg
 gccgcgctgggcaaggccacaatctgtacgcgtcgtagtagtacgcgcacgggcccagg
 ctacgctgcccgtggaccgtgcagctgggttaccgctacgcttggttaa

SEQ ID NO:46 polypeptide sequence of Orf23

25 LRQTTVPVRLVLRGAAVAQGDVVRAPETAPEKDGFGTTPVRPGLRVGLDQAPLELDVADG
 AQWHGATQSLDRLALGAGGQWRMSAASSVGELSMEPGAADVFGDAAGPGFQTLTVRTLGA
 AGSFEMRADAALAHADQLVVTDAQEGRHRVWLRAPAGAEPSKAQAVLVRAPADGKASFEL
 DGSDDRADFGTYRYGLAQPPGAWGLVRTGYSSTAAAALDTGGLGAVQGLWYAESNALGK
 RMGELRLNPDAGGAWGRAFSQRQRISPRAGRHFQQGVSGIELGADRAWFVAGGRWHAGWL
 30 LGYTRASRGFSGQKGHTDSVHVGGYATYIGANGVYADATLRASRFENSFDAPGWAGRTV
 SGSYRANGVGTLEAGRRLALDRHWFVEPQAEALAWFRAGGGTYTASNGLRIEDDGTSLO
 ARVGAQAGRRFDLRGGAUVQPYAQLSWVQELKGVSTVTRNGIAHRTDLGAGRVELGLGVA
 AALGKGHNLYASYEYAHGPRLSLPWTVQLGYRYAW

SEQ ID NO:47 polynucleotide sequence of Orf24

35 Ttgccttcgcccgcgaagaggcgccctcaggcggggcccgacgcgtccaagcagcggccg
 Gagggcctgcccgcgcccgatgccaatccccaacccgatgcaaagcccggagctgagatg
 Aaaccggccctgggggtggaacccggacctgaggcggaacctggtccgcaggggagcct
 gggcccccagcctggagcccggccgcaggacgagccgcacgcgcagccgctgcccgcggcc
 40 ggcaaccccgccgctgggatttacatgcccgcagcggcatcttgaccgcaccggttctg
 gccgtgctgggcacggccagtgccgcgcaaggtatctggcaggcagagatgaacgccctg
 agcaagcgcagtggtgaattgcggcttacgccggcagccggcggtggtgggcacgctcg
 ttgcgcgaacgcccaacgcctggacaatcaggtggtggacaggttcacccagaccgtgggc
 gggatcgagattggcgccgacacggccttgccggcgccgagggcgctggcatgtaggc
 45 gcggtggccggctacagccgtgcgcgcgcaagctggcgccacagcgcccgctggcaacagc
 gacagcctgcatgtgggcgcctatgcgacgtatcgggcgacggcggttctacctcgac
 gggattgtgcgggtgaaccgctacgagcacgatttcagggtgacggccagcgccggcg
 cgcgtgacgggcaagtatcgccgcaatggcatcgggctgtcgctggagaccggcaggcgt
 ttcacatgggcccggcactggttcgtggaaccgcaggtcgaagtggcggttcttccgttcg
 50 ggcggggcagactacacggccagcaatggcggtgcgcgtcgacgtggcaagcaccaagtgc
 ttgctgggcccggcaggcctgcaggtgggacgcaagctggatctgggcaacggcgaactg
 gtgcagccgtacgccaagctgagctggttcaggaggttcgatggcggtgggcaaggtgcgc
 accaagcatatcgcccatgacgtcaaactgcggggcgggcgccgcaactcgactgggc
 gtggccgcggcgcttggcaggcacagcagcgtgttgcgttcgtacgagtagcgaagggc
 55 agccgcttgaccattccgtggagctttcacgtcggctatcgatacgccctggttaa

SEQ ID NO:48 polypeptide sequence of Orf24

LPSPPEEAPQAGPDASKQRPEGLPAPDANPQPDAPGAEMKPRPGVEPGPEAEFPGPQGQP
 GPQPGARPQDEPHAQPLPPAGNPGAGIYMPSRGILTAPVLAVLGTASAPQGIWQAEMNAL
 SKRMGELRLTPAAGGVWARSFAQRQRLDNQVDRFTQTVGGIEIGADTALPAAEGRWHVG
 AVAGYSRARRKLAHSARGNSDSLHVAYATYIGDGGFYLDGIVRVNRYEHDFRADGQGA
 5 RVTGKYRANGIGLSLETGRRFTWAGDWFVEPQVEVALFRSGGADYTASNGVRVDVASTKS
 LLGRAGLQVGRKLDLGNGLVQPYAKLSWLQEFDGVGKVRTNDIGHDVKLRGGRAELDLG
 VAAALGRHSSLFASYEYSKGSRLTIPWSFHVGYRYAW

SEQ ID NO:49 polynucleotide sequence of Orf25

10 Atgagacgggttaaaggcccaggctttcgagggcagccgcagcaggccggcaggacatggg
 Gtggcgccctaccttgctggcgctggccctggggttccagggggcggcggcgtgggccaat
 Tgcacgacgtcaaacggtgctaccacttgaccaacgccaacggctctcataccaacaag
 Gtgggcagtgaggaccgagcgggatgaacgaacgcgtcaccgtgaatcagggggcgcgcatc
 GAgacaaacgccagcgcgggcgatcagtggtggaacgagcgggcaggtagaatcgagggc
 15 Ggtgcagtagtgcaaagcacgggtcaatactgctgctcggtccggccagtagcgcacaaacgctg
 Gaagcagcaagcaataacaatatcccatccaagtaaacgcgcagctcctggccaagggc
 Agcgcttcgcagtcacagcgcttgggattgtcaggcgccggcaataccgtcaccaaccat
 Ggcacgatccgggcccagataatgccgcggcaatctgggtcactgccaataccgccaatgcg
 Gccaataccatcgataactacgggactatcgaaacagtgctcaatggcggtacgccaac
 20 Gccatcggcagcagcggaacaacagtgccacggcgctggcgtagcggtagcgaatcat
 Gccaacggagcgcctcgccgcaacgtgaaagtcgaggctggcgacgacagcgctcatactc
 Gacggcggtctaccatcacgggatccttgaacgggtggcagcggcaacaacagcctgacg
 Ctgaagccggcgacggcagcgtgggcccgcgaatccgcaacttcggcacgatcaccaag
 Caggaggctggaacctggaccctgaatggccaggtcgggccgaacgacaacaacctcaag
 25 tccacggtcaaggtggaggcgggcacgctgggtcttcgcgggcgataacagcgggcgccacc
 caggcgcggtggttcgaggtgtccgcccggcgctacggcgagcgttaactgcccggcagcc
 atgcagtcacagcaacgcggcgacggttcagttcacgcaggacagcaatgccgctac
 gccggcgctgctgagcgggaccgggagcatcgtcaagcgcgggcgggcgacactgacgttg
 acgggcaacaacaccataaccggcaaggtggtggtggaggcgggcagcctcagcgtagtcg
 30 gcggccaacaacctgggtggcgaggttagtcggtacagctcaaggcgggcgccctcgcc
 ctcaagaaaacctcgtcgtcaatcgcgccctgacgctcgattccggggcgagacgttg
 atcatcgagccgggaacaaccacgacctggcaaggccaggttagcgggcgccggcaactg
 gtgacctcagggcgacgctggtgctggagcacgctccaatacgtatagcgggcggtacg
 gagatcaacaacggaacgctgcggggcgcgcatgatgccagcctgggttcgggcaggttg
 35 gcgctcaagaacagccagctggcgccacggacagcttcacggccacgcgtgcattgacg
 ctcgctggaacgaaagcatagacgtcgagccaccaagataactcagttggaacggcgaa
 atcagcggcgccggcaccctggtgaaggaaggccaggggaccttgctgctgcgcggaacc
 aatcagcaaaaatggcggcacgacctcaatgcggtagcgtgcagatatcccgcgacgcc
 aatcttggccgagggcgctggcgtaaacgacggcacgctgcagagcaccggcagcttc
 40 gcgacctcgcgcgcgccaccttgcgcgggccaggccaccatggaggctcgacgcttcgcat
 accgtgacctggaatggcgagctgagcggcgggcgcatgttcgcaagttagggcaggggc
 acgctggtcctggcgggcgccaacacgtactcgggtggcacgggtggtcgaggccggcgcg
 cttcgggcaggacacgaagacaacctgggacggggcgcaataacctgacggcgagat
 45 ctgcttgcggcgggcagtttttcgagcaaccgcatctcacgcttgctccgcggttccttg
 gacgtggctcgcgacgctaccctgacctggagcgggcgcatatcgggcgccggcgatctg
 gtcaagaaaggggacggggcgctgacactcacggcgctcaacgagtagcggcgccagacc
 gtgctccggggcggaagctgctgtggccaggagcgaacacctggggcgcgagcactg
 gtgctggaagacaataccgtgttcgagagcatgggctcgcatgccgccacgcggcaggtc
 acgctcaaggcgcgccccaaaggttagagacgcttgacggcactacgctcgaatggcgcgcc
 50 acggctcgacggcgacggcaagctgtacaagcaaggcgggcggaacgctcgtgctgagcggc
 aacaatacctacgccaaggcgctcgaggtctggggcggggtcgtgcaagtctctcgcgac
 cagaacctggggcgggccaatggcgcggtcacgctcaacggcgggcggttgggcgccaac
 ggggatttcaccagcaatcgccagctggagctgaccgcccggggccaaggccatcgacgtc
 gggcgggcaaggacgtgacgtggcgcggtgtcgtcaacggcgccggcgcgctgaccaag
 55 gccggcgacggtacgctggcgctggcgcgcccaacacctacaccggcgggcacgcgcttg
 caggcgggcaccgtgcaggtatcgcgcgacacaacctcgggcaggcgccggcgcggtc
 acgttcgacggcgggcggtggccaacacggggcagcttttcgacccgcgcgacggccacg
 ctcaacaaggctggccagatcgataccgacggggcaccacgctgacatggaacggcgcc
 atcgggcggaaggcgagctgcgcaagcaaggggcgggcaccctggtgctggggcgagcc

aacacttaccagggcgacaccccgctcgaggccggcacgctgcaagtgtcgggcgacgcc
 aatctggggccagggcgccgtgcatctgcacgacagccggctggcgacgacgggtacgttc
 gcgacctcgcgccgtctggagttgaccggacgtggcgcggtgcaagcggtcgcccgcc
 5 acgctggattggcgcgggacgggtcgctggcgccggcacgctgggtcaaggagggcgaggc
 acgctgggtgctggccgggtgacaaccagcatgcccggcgccacggaggtcagggccggcacg
 ctgcaggtatcgcgcgccaccaacctggggcccgcgctggcgctggagaacgcggcg
 ctggccacgacccgagcttcacggccacgcaggcagccacctgactggcaacgccgcc
 atcgacacggcccgccggcaccacgctgggatgggagggggccatcgggcgaaccggcagc
 10 ctgcacaaaaaggcgagggaagctgggtgctgggtcaaggacaaccacatgacggcggc
 accacgatccacgcccgtaccctgcaggtgtcgcgcgacgccaacctgggctcgggacag
 agcgcggtgacgctggatggcgcgccctggcggtttctgccgggttctccagcgggcg
 gagatcgtcgtggcgccgggacggcgcgcttctgggtgacggcgggccacacctgcaa
 tggcgaggccaggtcgcgggggcgggcggttgaccaagacggcgagggcgacgctcgtg
 ctgacacctgaacaacagccgctgacgacccgggtaccgggttaccggcggggtgctgcgctc
 15 tcgcgcgatgagaacctggcgaggcgcatggcatgctgacgctcgacggcggcacgctg
 tcgaccaccgcccgggttcgcgagccggcgcaacgccacctggggcaacggcgcgccgg
 atcgtcgtcgccgacgcccacgctggatttgacggggcgacgttgccggcgcgggccgg
 ctggtcaaagagggcgcgggcgacgctggccttggggcgccacgaacacctatgcccggggc
 accgtggtcgaggccggcacgctgcgggtcgcgcgcgacgccaacctggggcgggcgcg
 20 ctgacctgaacaacagccgctgacgacccggcgccggcttggccacggccgacggtg
 acctctccggggcgcgccctcgatcgacacccgacgacccggcgacgctgcaatggcgcggc
 acggtcaatggcgccggcaggctgggtcaagcagggcctgggcacctgggtactggacggc
 gacaaccggtacggggaggcaccgaggtcaatgccggcacgctgcaggtcgcgcgcgac
 gccaaacctggggcgggcgacgtggcgctcaatggcagcagcctggccgacgacggccagc
 25 ttgcaccgcgcgacggccacgctgagcgcgcgcccgccatcgacacggccgacggc
 gccaccttggaactggatggcctgctgacgggtgacggcgccctgggtcaagcagggcaac
 ggcacctggcgctggcccgggccaacctgctatggcgggcgccacctcgctcaaggcgggc
 gccgtgcggatcgccccgacgccaacctggggcgggccggcaccggcgtaacgctggac
 ggcggcgcgctggccaccacggcggtatctcgcgacggcgccgcgcgggcgacacctggggcgcg
 30 gccaaacggcacgctggacgtggccgcccggcaccgcctggactggaacggggcgatcgcc
 ggcgcccggcgctgaccaagacggcgcccgccacctggcgctcaaccacgacaacacag
 catgccggcgccacctgggtccatggcgccacgctgcggtatcgcccgacgccaacctg
 ggcgcccggggcacggcggtgacgctggacggcgccacgctggccaccacggcgctcggtg
 ggcggccgagcgcgcgctgcgctcgggcgcgcaacggcggtatgctgcccggacgcccggc
 35 acgacctgggttggcgggcggtgggtcgccggcgcggggaagctgaccaaggccgggtccg
 ggcacgctgggtgctcagcgccgataaccggcatggcgggcgccacggcagtcacggcggt
 acgctgcaagtctcgcgcgacgccaacctggggcgcgggcgccggcgccctgacgctggac
 ggcgccaccttgctgagcaccgcccagctttgcctcgggcgctgtgcccacctcgatgcc
 gcgggcgggcaccttcgtcaccgcgacggcacccggctggattgggacggcgcgataggc
 40 gggggcggtggcctgggtcaagggaaggcgccggcgagctgcggcttggcaatgccaatacc
 taccaggggcccagcccgcacgcccggcgccgctggcgctcaacggcagcatcgccagc
 ccggtcacggtcgagcaggccggcggtgctggggcgccacggggcgcatcgctcggggatgtg
 gccaaacggcgcggtgggtcgcgccgggcaactcgatcgggcggttgacgggtggccggcaat
 45 tacgctgggtacggggcgccagcctggaagtggaggcggtgcttggcgggcgacggcgcccg
 gccgatcggtgggtgctcgacggcgcgccggccagcgcgctcacgcccgtcgtagtcaag
 ccgagggcggggtggggcgccctgacctgcggcgcatcccggtgggtcggtggcccagggt
 ggcgccacgaccgcgcccggggccttcgcctggcgacggcgctgggtcgcgggcgccctac
 gagtaccagtgtgctgcggcgcgggcgacggcgcccgcgcgagggcgcaagactgggtac
 ctgctacctcccgcgtcgagcgcgacaaggcgggcgaggatcgtaagggtcggtgcccttc
 50 taccggccccgaggtggcgctgtatgcggcgacgcccacgctgatgcgcatgggtcggcacg
 gaagcgctgggcagctaccggaacgcgcgggccagcccggcgcgcccgccgagggca
 ggcgcccagcccggcggtggcggtggggcacgtaccttcggggcgctggttcgagcgctcc
 gcgggcagcgaagcgggcgccgctcctcaacggcagcctggccggcatgcagctggggcgcg
 gacctctacacgctcgctcgggcaccggcgatgccgacgcttcggcggtgttcggcgga
 55 tacgctacggcccgccggcgatgtgcggggcgccgctggcgcgcgcgagatccaggcggtgggc
 acgtccacgctgcgggcgcccgagctggggcgccactggacgcacactgggtccgagcggc
 tggtagctcgacacgggtgctggcgggcgacgctgacaagcagcagaccagctcgctggcc
 catgtcgggcgacaccagccggcggtggggcatgatggcctcggtggaggccgggtaccgg
 tggcagctcaatccgcgctggcaaatcgagcccgaggccagttgggtgtatcagcagctt
 60 ggcacgccaatggcgccgacccgctgtcttcgggtgctgtaagacgcccgatgcgctg

acggggcggtggtgggcacgcgcctggcgggccagtagcgcatacgggaaggcgagttgcgg
ccgttccatggggtatcgctgctgcacgatttcacggcgccgacaccgtcacgttcgcg
ggcgtgcacagcgtacgcgccagccgccagaacacggccgtggatctgaaggcgggcggtg
gacacgcagctgggcaagagcgtaggcctgtgggggaggtcggtacggcaagtcggtc
5 ggacgcggcgacggcagcgaccgtggctggagcgccaacctggggctgcgcgtggcgat
tga

SEQ ID NO:50 polypeptide sequence of Orf25

10 MRRLKAQAFEGSRSRPAGHGVAPTLLALALGFQGAANCTTSNGATTCTNANGSHTNK
VGSGPSGMNERVTVNQGARIETNASAAISVGTSGQVRIEGGAVVQSTVNTAASGQYAKTL
EAASNNNISIQVNAQLLAKGSASQSSALGLSGAGNTVTNHGTIRADNAAAIWVTANTANA
ANTIDNYGTIETVLNGGYANAIGSTRNNSATGAGVTVRNHANGRIVGNVKFEAGDDSVIL
DGGSTITGSLNGSGNNSLTLKAGDGTLGRAIRNFGTITKQEAGTWTLNGQVGRNDNNLK
15 STVKVEGGTLVLRGDNSGATQGGVLQVSAGATADVTAASAMQSI SNAGTVQFTQDSNAAY
AGVLSGTGSIVKRGGGDLTLTGNNTHTGKVVEAGSLSVSAANNLGGAGSSVQLKGGALA
LKKTIVVNRGLTLDGSAQTLIIIEPGTTTTWQGVVSGAGKLVTTQGGTLVLEHASNTYSGGT
EINNGTLRAAHDAISLGSGLTALKNSQLAATDSFTATRALTLAGNESIDVAATKILSWNGE
ISGAGTLVKEGQGTLLLRGTNQNGGTTVNAGTLQISRANLGRGALALNDGTLQSTGFSF
20 ATSRAGTLRGQATMEVDASHTVTWNGLSGGMLRKSGQGTTLVLAGANTYSGGTVEAGA
LRAGHEDNLGRGAILTQGGDLLAGGSFSSNRDLTLVRGSLDVARDATLTWSGAISGAGDL
VKKGDGRLTLTGVEYAGQTVLRGGKLRVARDENLGRGALVLEDNTVFESMGSHAATRQV
TLKGAPKVFETLDGTTLEWRGTVDGDGKLYKQGGTTLVLSGNNTYAKGVEVWGGVVQVSRD
QNLGAANGAVTLNNGGLAANGDFTSNRQLELTAGAKAIDVAAGKDVTRGVVNGAGALTK
AGDGTALAGANTYTGTRLQGGTVQVSRDNNLQQAAGAVTFDGGRLANTGFSFATARTAT
25 LNKAGQIDTDRGTTLTWNGAIGGKGLRKQAGTLVLGGANTYQGDTRVEAGTLQVSADA
NLQGGAVHLHDSRLATTGTFATSRRLLELTGRGAVQAAAAATLDWRGTVAGAGTLVKEGAG
TLVLAGDNQHAGGTEVRAGTLQVSRATNLGPGALALENAALATTASFTATQAATLTGNAA
IDTAAGTTLGWEGAIGGTGSLHKKGEGLVLDKDNHHDGGTTIHAGTLQVSRDANLGSQ
SAVTLDGGLAVSAGFSGREIVVGAGHGALSVTGGHTLQWQGVGGAGALTKTGDTLV
30 LEHDNTHAGGTRITGGVLRVSRDENLGEAHGMLTLDGGTLSTTAGFASRRNATVGNNGGR
IVVADAATLQGDVAGAGRLVKEGAGTLALGGTNTYAGGTVEAGTLRVARDANLGGGA
LTLNNSRLHATAGFATGRDATLSGRASIDTDDRATLQWRGTVNGAGRLVKQGLGTLVLDG
DNRYAGGTEVNAGTLQVARDANLQAGDVALNGSSLAATASFATARTATLSGAAIDTADG
ATLDWNGLLDGDGALVKQNGTTLALAAANRYGGGTIVKAGAVRIARDANLGRAGTGVTLD
35 GGALATTADLATGRAATLGAANGTLDVAAGTRLDWNGAIGGAGALTKTGAGTLALNHDNQ
HAGGTLVHGGTLRIARDANLGAAGTAVTLDGGTLATTASLAPERALRVGARNVLLPDAG
TTLDWRGVVAGAGKLTAKGPGTLVLSADNRHGGGTAVTGGLQVSRDANLGAAGALTL
GGTLLSTASFASARVATLDAAGGTFVTRDGTLDWDGAIGGAGGLVKEGAGELRLGNANT
YQGPTRIAAGRLAVNGSIA SPVTVQAGVLGGTGRIVGDVANRGVVAPGNSIGALTVAGN
40 YAGTGGSLVEAVLGGDAAPADRLVLDGGAASGVTPVVVKPQGGVGGTLRGIPVVVAQ
GATTAPGAFLAQLVAGAYEYQLLRGAGDGAQAQDWYLRISRVERDKAGRIVKVVPF
YRPEVALYAGTPMLMRMVGTALGSYRERAGQPGAAPEAGAAARRGVWARTFGRRFERS
AGSEAAPSFNGSLAGMQLGADLYTRRSATRHADAFGVFGGYATARGDVRGLARGEIQAVG
TSTLRAAQLGAYWHTGTPSGWYVDTVLAGTRYKQQTSSSAHVGATSRGWMMASVEAGYP
45 WQLNPRWQIEPQAQLVYQQLGIANGADRVSSVSYPKTPDALTGRLGTRLAQYAYGKAQLR
PFMGVSLLDHFTGADTVTFAGVHSVRASRQNTAVDLKAGVDTQLGKSVGLWGQVGYGKSV
GSGDGS DRGWSANLGLRVAY

SEQ ID NO:51 polynucleotide sequence of Orf26

50 Atgttgccaactggagccccgatgcgtagcgcgcccgccgctacccagcagcagcttgcc
Ccggtgcccggccatgctggccgcccgcgctgcaatccttggcgcgacgcccgcg
Gccgcccctgtgcccacgctggcgccagccagcgggcagagcgtgcaatgcgacggc
Gccgtgggtcaaccagtcggtcgaggcgccgcccagccagacgtaaccatcacggtg
Gcgcccggcgcgctgttctcgaccaacgcgacgcccgcgctgtcggtcgatgaccgcagc
55 Cgtatcggtcaacgagggcacgatccagatggccggcgccgcccgcctcgcgccgccc
Atgggtgggtttcgccgacaacaaccagttgatcaaccgcggtcgatcacgacatcgggc
Agcgccgtgcgcggcatctcggtgcccaacgctcgccagtagtaccgggacgctgggtcgacaac
Agcgccagcatccgcacccagggcgccagcgccagggcatcgccatcaacggcccgccc

Aaccgcgtccagaacagcggcgccatcacgcgtcaacggcaccgacgccaagggcggtgat
 Ctgaaggcggcagcccgcgcccaacgtactggtcaacggcggcaccatccatgcgcgc
 Ggcgcagcagcaacggcatattcgccccgacggcggtgcatgtgaacaccaccaacgcc
 Aacggctttcatgcccgcgtcgagaacctgcccggcgggcgcatcctcagcgatcactcg
 5 Tatgcgctgcgcgggcagaacggcaacgataccttcatcaacgcccggctacctgcaaggg
 Caccggcgggcgccgcgcgacacggccgtctacatgggccccagggcacgggcacgctg
 Atcctgcgcaccggctcgcccatcgccggcctggccgacggcgggcgggcgccagccat
 Gcgtatctggagggcagcggcacgggtggacaacgggttcgccaacttcgcaccctgacc
 Atgcgcggcgccgactggcgctggacctcgagcgccggttcaccgaaagcggtggacctg
 10 Cgcaccggcacattctttcttgcggcacgctcgccagccccggccaaccgcctggccgccc
 Ggcggggtgctggccggcaccggcacgctggccggcgcatcgcaacggcgcgaaatc
 Cggcccgccccgaacgacggcagcggtacggcgcgctgacgggtgcgcggcgattacacc
 Ggcggggggcgcgctgcgcgctcaacacgggtgctggccggcgacggggcgccctcgga
 Aggctggtcatcgatggcgggcacggcgccgagcagcaccgggtcacgggtggtcaaccgg
 15 Ggcgggcagggcgcgctgaccggcgccgacggcatcctggtggtccaggccatcaatggc
 Gcaagctcggacgcggcgcccttctcgctggccgccccctcaacgcccggcgcatacgag
 Tacaggctgtaccggcgggcgccacggggcgccgcccggacagctggtacctgcgctcg
 Cgggcctatctggtcgaggaccaactggccggcagccttgccgaagccgaggcgatcgcc
 Gacgacatcgccggcgccacggcgagcgccgagcatcgaggacacggcgctgtaccgg
 20 Cccgaggtggcgctgtacagcagcattcccatgctggcgcgccgcatgggcctggcccag
 Ctgggcaccttccacgaacggcagggcaaccaggcgctgctggcgcgcgacggcgaaacgc
 Gtcgcgccctggcgcgcgccatggcgggcaacagcaagcaggcgctcgacggcgatgcg
 Caaccggcatcgacgcccgcctggccggcggtgcaactggggcaggacctctacagcagc
 Gtcgccccggacggcggaacagcaccgcttcgggctgttcggcggtatggccaggcgcg
 25 Ggcgacaccacggctcgccggcgggcgagcgcgacggcgctaccggccggctgaccatc
 Gacggctacagcgctcgccggctactggacctatgtcgcccgcgcggggtggtacgtggat
 Gccgtgctggccaaacacctggatggacatcgacaccgactccaaggccggcgcgacgcc
 Gatacgcgggccaggcggttcacggcttcgctggaaagtgggtaccggctggcgctgtcc
 Gagcgctggacgctggagccgacggcgagctcatctaccagcacacgcgcgtcgacgg
 30 Ttctcgacgcgctgtccgaggtgcgcatccgcgacgacaacgcgctgaccgcccgcctg
 Ggcgccccggctgcagggcgagtagcgccggcgccgagcgaggtgtggcgccccctacgcgcg
 Ctgaatttctggcgcaccttcagcgcgagaaacaccgtcgctgctggcggaagacagcatc
 Gatacccgggcgggcgacctcgctcgaaactggcgggcgccagcggtgacgctggcc
 Cgcagcctggccctctacggcaggctggcctatgccaccagcatcgacagccagtatctg
 35 Cgcggcgcttcggcgcgagctggggatgcgctacacctggtaa

SEQ ID NO:52 polypeptide sequence of Orf26

MLRTGAPMRSARRRTPARLAPLPAMLAAGLLQSLGATPAAAACAPTLAPASGQSVQCDG
 AVVNQSVAAAAGSQNVITVAPGALFSTNATRALSVDDRSRIVNEGTIQAGGAGASRGA
 40 MVGFGDNNQLINRGSITTSGSGVRGISVPNVGSTGTLVDNSGSIQTQASAHGIAINGPG
 NRVQNSGAI TVNGTDAGVYLQGGSPAANVLVNGGTI HARGASSNGIFGPDGVHVNTTNA
 NGFHARVENLPGGRI LSDHSYALRGQNGNDTFINAGYLQHGAGGRDTAVYMGPGTGTTL
 ILRTGSAIAGLADGGGAASHAYLESGTVDNRFANFRTL TMRGADWRWTS DAAFTESVDL
 45 RTGTFFLAGTLAS PANRLAAGAVLAGTTLAGALRNAGEIRPGPNDSGYGALTVRGDYT
 GAGGALRVNTVL AGDGAASDRLVIDGGHAGGSTPVTVVNRGGQALTAADGILVVQAING
 ASSDAGAFSLAAPLNAGAYEYRLYRGATGAAPDSWYLRSAYLVEDQLAGSLAEAEAIA
 DDIGRRTGERPSIEDTPLYRPEVALYSSIPMLARRMGLAQLGTFHERQGNQALLARDGER
 VAAWARAYGGNSKQALDGAQPGIDARLAGVQLGQDLYSSVRPDGGQHRFGLFGGYGOAR
 GDTHGSAGGERDAATGRLTIDGYSVGGYWTYVGPRGWYVDAVLANTWMDIDTDSKAGRDA
 50 DTRGQAF TASLESYPLALSERWTL EPOAQLIYQHTRVDGFS DAVSEVRIRDDNAL TARL
 GARLQGEYAAAAQVWRPYAALNFWRTFSGENTVVLGEDSIDTRRGATSL ELAAGASVTLA
 RSLALYGRLAYATSIDSQYLRGASQLGMRYTW

SEQ ID NO:53 polynucleotide sequence of Orf27

ATGCCTGCTCAACGCACTCCTCGCACCGCGGTTTGTGAGGCCACCGTCCGTTTCATCGCCA
 CGTTGGATCCATTGCACCGGATTCTGTGCTGTGCGCGCTGCTGGCGGCATGCGGAGGCGGT
 GCGCGCGGAGGTGGCGGCGGAGGCGGTGGCGGCAGCCCCGGGCGGCCGTGCGCCATCGGCG
 CCGCAACCCGCCCCCTTCGCCACGCCCCGAACCTGCACCCGAGCCGGCACCCAATCCTGCG
 55 CCCAGGCCTGCTCCGCAACCGCCGGCGCCCGCGCCTGGAGCGCCCCGTCCTCCCGCGCCG

5

10

15

20

25

30

3

4

4

5

:

HVDDIRPGYVGGDGKSVFVIKAGAVSGVFATLTRSPGLLLNARLDYRPQAVYLTMRRAER
 VHAAQRGADDGRRASVLAVAERLDAAMRELDALPESQRDAAAPAAAIGRIQRVQSRKVL
 QDNLYSLAGATYANAAAVNTLEQNRWMDRLLENHLAQAGGERVAAIAEYRHGQLRWRPDGL
 QGRQRGNGIMLGLAREVSAGLSLAAALTHSRTHWDESSGAPARDNAAMTTPGVLLGARRA
 5 WEDGWVQVQALGYSRYRNQATRHISLGDAGHTVGATARGHVWQADAGLGRQWTLAPGHITL
 APRAGLQLTHLRQQGFSESGAQGLGLRAHALTRTVPTLWAQLQSRHAFMLGATPMTAQLQ
 LGVWHDLRARRYAASGGFAGLAQDQASGYWVPVPRTRVQALGLRAEFAPGLVLGLGYTG
 QLATHWVDHQLSASLTTRY

10 **SEQ ID NO:55 polynucleotide sequence of Orf28**

Atgtcgtccccgcgtccccccgcaccttggcgcgccaccgctcgcgcttgccggcctgtcg
 Cttgggttgcgctgccggcgcatacggcgcgcccgcgccggcacaaaccgctcgtcaccctg
 Cccgcgcaagaggtcatcgcgacagcgctcgcgcgcccggtccgtgctgcgccctgccg
 gagatcgagcgcgcgaggccgacaacttcgcctccctggatcgatcagctgcccgggcatc
 15 tcgatggccggctctccgcgccccggcgggcaagcctgaacatctggggcatggggcat
 accgaggacgtgaaaatcgctcctcgatggcgcgcccaagggttctcgagaagtaccggcag
 ggctcggtcttcatcgaaccggaactgatccggcgcatcgaggtcgacaagggggcgac
 aacctggctcgacggcaatggcggggttcggcgccaccgtcaagatcgataccaaggatgcg
 gccgacctgttgccgcggggcgcgcgcttcggcgcgctggccaagtacggccgccattcg
 20 aacgacggccaggacatctacagcggtggcgctgtacggccgcaccgcgcccagcggggcc
 gacggcctgtgtatgccaaccgcccgcgacggcgggcgatctgcgcgcccccgacggcacc
 cgcttcgcatactcgcgcaacaaccagcgctcgctgctcgccaagggtcaacctctatccg
 gacgagcccccagaccatcacctgtcgggccatgcgttcgaatgcggcggggttggaaccc
 ttgcgggccaagcgcgacgatcttcccgcgccttcgcaggccgatatcgaccgctacggc
 25 ctgaccgaagcatggcgggcgcaagctggtccatcgcgaccagctcgaccagaactacagc
 gcgaaatggaacatcgccccatccgccatccctgggtgaacctcacgctggcctatgcc
 cgctcggacacccggcagcgcgaccggcgctcgctcccggcgctcgagtcggccttctc
 ggcacgctgggcaacaagagtgggtcgactaccgcgacgacgggttcgacctcagcaac
 gaaagccacgtggccctgggcacggccgagcatgtcctgctggcgggcctgcgctggcac
 30 cggcatcgccgcgacacgctcatgtactacccgcccggcgcgggcgagcccgattacaac
 cacgggtacttccagccgcactacatgccttcgggcacgcagaccgtgcgcagcctgtac
 ctgcaggacgccgtcacgctcgggcgcccttacgctcacgcccggcgctgcggtacgacct
 gtgcgaataaccggcaggccaaacgacgcgccccgctacaacaacccccgcccccgggc
 gggcatgactaccgcgcgctcgtacgcgggctggaccccgacactgggctgggtctgg
 35 aaggcgcgcgaggcggtggcgctgttcgcccgcgcccggcgccacactggcgcgacccgctc
 atcgacgaacagtacgaagtgaatatgcgaagtcgaatgtgtcgggcagcagccgggcg
 ctgcccggcagcgcatcggtgggcctgcgcgcccggcgccgtaactggattacaacgatatc
 gcgacgcgcccgcagacgctgcagatacggaccacgctgttcgcaatcgcgggcaagcac
 gagatcttccagccgctggcggtggcatgcgcgggcgaggccgagggcgcgccgctcg
 40 gactggcccaagccctgtccaactaccgcaacctgcccggctacaccatcgaagggctg
 gaactggagacctactacgacagcccggcgatgttcgccagcctgtcgctttcgcccatg
 cgccggcaccgcgacgcctcgccgcgcgatccatggggcgcgccacctggatcgccgag
 atcccgcgggtctcgccgcgcgcgatgctggcggtgaaactgccgcgcctggacatggctc
 ctgggatggcgcgggcgaattcgctgcgcgccaggaccgctcgccgaccgacggcgacccg
 45 ctggccggctactgggccttggccaaagaccgcccggctacgcgctgcacggcctgttcgca
 agctggcaacccccggcatgtcaaaaggcctggacgtgcgcctggccgcccgaacacctctc
 aaccggccctatcatccctacctgggcgaagcggatcgggcacggggcgcaacatcaag
 ctgagcatcgcccagcgcttctag

50 **SEQ ID NO:56 polypeptide sequence of Orf28**

MSSPRFPAPWRAPLALAGLSLGAAGAYGAPAPAQTVVTLPAQEVIGDSVAAARSVLRPL
 EIERAQADNFASLVDQLPGISMAGSPRGGQSLNIWGMGDTEDVKIVLDGAPKGFKEYRQ
 GSVFIEPELIRRIEVDKGFPHNLVDGNGGFGGTVKIDTKDAADLLPPGARFGALAKYGRHS
 NDQQDIYSVALYGRTRADGADGLLYANRRDGGDLRRPDGTRFAYSRNNQSRSLAKVNLYP
 55 DDAQTITLSAMRSNAAGWQPFPAKRDDLPAQSADIDRYGLTEAWRRKLVRDQLDQNY
 AKWNIAPSAHPWVNLTLAYARSDTRQRDRSSRASQSAFLGTLGNKSWVDYRDDRFDLN
 ESHVALGTAHEVLLAGLRWHRHRRDTLMYYPPGRGEPDYNHGYFQPHYMPSTGTQTVRSLY
 LQDAVTVGGTLVTPGVRYDHVANTGRPNDAPRYNNPAPVAGHDYRRVSYAGWTPHLGVVW
 KAARGVALFADAGRTWRAPVIDEQYEVQYAKSNVSGSSRALRPERIVGLRAGAVLDYNDI

ATRGDSVQIRTTLFRNRGKHEIFQRRGVACRGQAEGGAASDCPKPLSNYRNLPGYTIEGL
 ELETYYDSPAMFASLSLSAMRGHRDASPRDPWGPRTWIAEIPPVVSARAMLGVKLPRLDMV
 LGWRGEFVRRQDRSPTDGDPLAGYWALPKTAGYALHGLFASWQPRHVKGLDVRLAADNLF
 NRPYHPYLGEAVSGTGRNIKLSIAQRF

5

SEQ ID NO:57 polynucleotide sequence of Orf29

Atgaaggcgcggcgccctggccatggcgggtctgtcgctggcgctcgggcggtgctcgctg
 Tcgagcagatgcaggccatgcgcgacgccgcgacgtccctgcgcgcacgcctgctcgaa
 Gggcagcaggccgtgggcccggcgccgagcggcgccggcgcggaagccgccaggacgtc
 10 gcgcggccctggctggccggcgcccgccagccgctggcacgcgaggtgctgctgccgcg
 gcgctgcgcgcgatgctcgatacgaccctgctgttcgcgggcaaggccacgctgcccg
 ctggccgagcgccctgcatcgcgccaccggcatcgccgtgcgcgtgcatcccgcgcgctg
 ctgcccgcgcgcgccttccctgccgcgcctggcggggagggccgagctggcgccgagcct
 cccgccaggccgaactgcggggccggcgccgctcgctggccgacacgctcgacgcgctg
 15 gccgcgcagctgtacgtgcaactggcgctaccatcgcgggcgccatcgagttctaccgcacc
 gaaacgcgggtcttcgatgtgcgacgctggcgctggccgcgagcgcgaggtcggtcg
 ggccgcgcggcgagcgagacggcgagtttcgaccatgcctcgagcacggtgctcagc
 gccgacgccggcaaggcgctgcaggccgtgcgggaccgcgctcgccgcttccctgacgcgc
 gccggcgctcatcgccgagatcgaggcgggcggaagcacgctcgcggtcacggatacgccg
 20 gaggcgctcgcgcgcatcgaaaaatacctgcaaggcgagaaccgcgcctcgacgcgcgg
 gtacgcctggtgttcgaagagctcacggtgcgccaccacggccgcgcggaaggcggtc
 gatggcaggcggtctacgccagcgcgccgcgcgcggcgctcgtaacgcatgcccgcggg
 gccggcgccggcaggcgcgctcgggggcccgctgctggcgggccctggcgcgacgcgcg
 gccctgatcgccgcgctgagcaccatgggagcggtactgcgccatcgagcatacccatg
 25 ctgacgctgaaccggcgcgccgtcacccacgcgctgcgcaccacgttttccctacgtggac
 caggtgcagcgccctgagcccgaccgcggcgcccggtggcgcgcatgccgtgcccggg
 ctggcggtgcagcagaagcgcgagacgggtgggcacgttccctacgctgttgcccaggcg
 cgcatgacggcgcgcatcctgctctccatttccctatgacaacaccattgcccgccgctg
 cgcaccctgaccttcggcgagggcgccagcaagtgtcgctgcagcagatcgccatcgac
 30 ggcagcgccatcgctgcagcaggtcgagctgctgcccggccagcccgctcatcctgtcgggc
 ttcgaccacagcgaagaccaatacgaacgccaccgcctgttcccgatgcgcgcgctcgcg
 gccggcgggcgacgaccgcacggcgcgcgagcggggtcacgaccgtggtcatggtcacccgcg
 cagatcgacgaggggtga

SEQ ID NO:58 polypeptide sequence of Orf29

MKARRLAMAGLSLALGGCSLSQOMQAMRDAATSLRARLLEGQQAVGRAGERPAREAAQDV
 ARPWLAGRAQPLAREVLLPPALRADVDTTLLFAGKATLPVLAERLHRATGIAVRVHPDAL
 LPRAAFLPRLAGQAELEAEPQAQELRAGPRPLADTLDALAAQLYVHWRYHRGAIEFYRT
 ETRVFDVRTLALAASAQARLGRAGSGTGSFDHASSTVLSADAGKALQAVRDRVA AFLTR
 40 AGVIAEIEAGGSTLAVTDTPEALARIEKYLQGENRALTRRVRLVFEELTVRTTAAAEGBI
 DWQAVYASARAAASYAMPGGAGAAGALGARVLGAPWRDARALIAALSTMGA VLRHRSIPM
 LTLNRRAVTHAVRTTFSYVDQVQRLSPTAAAPGGRDAVPGLAVQQKRETVGTFLTLPEA
 RDDGRILLSISYDNTIAQPLRTLTFGEGGQQVSLQQIAIDGSGIVQQVELLPQPVILSG
 FDHSEDQYERHRLFPDAPLAAGGHDR TARERTTVVMVTAQIDEG

45

SEQ ID NO:59 polynucleotide sequence of Orf30

Gtgaccatgttcatccgctggctcattctctccgcctgctgctgctggccgcctgcagc
 Cgcgctcccgataccgagatcctgcagcgcgatgtcgccagaccctggccgccacgtac
 Ggcccggacctgttcgacatcgctcgcgctgcgcgcgcatgggctcgccaccgacagcagc
 50 gcccgcgcgggcccagacgcgcgggtggtctattacgatgtggtgctgggctgaagaag
 gacctcaccctgggcgccctgggaccagcccggcgccgcccgcgctgggtcagcctgctgggc
 gccggggcgcgcgagcatctcgggggtgaaatccagcggcaatgccgcggcgaccagatc
 gtcgcccacgccagcgccatctaccagcgcgacgcagagcaatgggtgcacgtcgccccg
 gccagtttcagggccaccgaagcgccctcgctggacaccggcgcgccgcgcggcggtgacg
 55 cgccagctgctccagacgctggagcagatcacgcgttccgtgccctacagcgccctcagc
 accgcccagcacgtggtgcaacaggagctggagcgctcggtggcgcgcatcaatggccgg
 cttgcccgcctgcaaaagggtaccgcgtggcgaccggccccgacaaggcgagtaacctg
 gcgttcggccaggcgctggccgcgcatcgggcgcaacgagcaggtgcgcgctcattccctc
 attaccggcgccgagcgaacacatggcatgctgcgcagcggcgcgggcggtggccgccc

10

20

25

30

3.

4

130

MTLFDGGARSARVEQARAAYDAQAAAYRQSVLTALREVEDYLVQLRVMEHEQQVQRNALE
SARESLRLARNQYEQGLIDYLSVAVLETTALNTERNAISLLGSRLNASVQLIAALGGGWQ
GLPAEAAASAAAEPSAP

5 **SEQ ID NO:63 polynucleotide sequence of Orf32**

Atgacgcatcccgtcccagacgacctttgcacgtaccgcccggcgcgctgcttgcgcgctg
Gcgctggccgggtgcgcccgtggggccgcagtagccaggcgcccacgcccgcgcccgtgaag
Ctggccagccccgaacaggcgctgttctcggccgacgggttgcaacgcgaatgggtggcgc
cagttgcaggatgcccggctggacgcgttgatcggcctggcgctggcgcgcaacctcgat
10 atcggcctggcgctggcgcgcaacctcgatatccgccaggcgaggcgcgccctgcgcgaa
gcgcgcgcccgcgtcgacgaaaaggaactggaccgctggccgaccgtgaccgcggcgcc
ggctacacgcgcagcctgtcgagatcaacccccggccccgaccagcgcaacctcgcgcaa
agctaccgcgcgggcttcgacgcgacctgggaaatcgatttggtcggccgcctgcagcga
cggggcggaggcgcgccgcccgcgcgcgaccaggccgcgcgcgcgacctggcccagacgcgc
15 ctgggtgggtggcgagctggcacgcaactatttcgagatgcgcggcgccgagcaacgg
ctggccgtggcgcgcccaacctcgccaccagcaggagacgctgcgcgtcaccgcggcg
ctgggtggaacggcgccgcggtatgcggcgacctggccagcgcaacggggcgagctggcc
ggcacgcggcgctgctgcgcgcgtggagacgcaacggcgccctggcccagtagcacatc
gccgtcctggcgcccatgcggccggccgagctggcgagctgcggcaggagcagccgctg
20 gcgcgcgtggccgcgcaattgcccatcggcgacctggccatgctgctgcaacgcccgc
gacgtgcgcgcgcgcgagcgctgctggccgcccacaaacgcgcgacctgcggcccatcacc
gccgaactgtatccgcgcacatcgacctggcggggttcctcggtttcattgccttgcgcggc
ggcgacctgggcccaggccagcagcaaggccttcgcgcgtggcgccgacgatcagctggccg
gcgttgcaacctgggcagcgtccaggcgagctgcgcgcggggccaggcccgcgcgcgcg
25 gcgcggggcgcgctacgaacagggtggcgctgcaggccatcgaggaagtggaggcgcgctg
acgcgctatggacagaaccagcagcggtgcgcgcacctgcttgacagcgccacgcagagc
cagcgcgccgcccagctggcgcaaacgcgcctatcgtgaagggggcgcgccgtatttgacg
gtgctggagcgcgagcgtactcttttcgcgcgcacaggatgccgtggcgcaatccgagtcg
gagtcctataccagcctggctgcgcgtctacaaggccctgggcggaggctggaataccgac
30 gccgcccgcgcccgcgcttcgcgcgcgaccgcgcgcctgcgggccagcccctga

SEQ ID NO:64 polypeptide sequence of Orf32

MTHPVPTTFARTAGALLAALAGCAVGPOQAPTPAPVKLASPEQALFSADRLQREWWR
QLQDARLDALIGLALARNLDIGLALARNLDIRQAQARLREARAALDEKELDRWPTVTAAG
35 GYTRSLSQINPGPDQRNLAQSYRAGFDATWEIDLFGRLQRRRAEAAAARDQAAAADLAQTR
LVVVAEELARNYFEMRGAEQRLAVARANLATQOETLRVTAALVETGRGYAGDLASARAELA
GTRALLAPLETQRRLAQYHIAVLAAMRPAELGELRQEQPLAPLAAQLPIGDVAMLLQRRP
DVRAAERLLAATNADVGAI TAELYPRIDLGGFLGFIALRGGDLGQASSKAFALAPTISWP
ALHLGSVQAQLRAGQARHDAARARYEQVALQAI EEVEGALTRYGQNQQRLRDLDSATQS
40 QRAADLAQTRYREGAAPYLTVLDAQRTLLRAQDAVAQSESESYTSLVALYKALGGGWNTD
AAAPARSARTALPASP

SEQ ID NO:65 polynucleotide sequence of Orf33

Atgaaacctgtcgtcatgagaaccttggtgtcccttgccgtggccacggccctggccggc
45 Tgctcgctggcgcccacctacgagcgcccgcaggcgccggtcgacgcggcctatccgtcc
Ggcccggcctacggcgcgccggggccaggccgcgcggcgcgccggccgcgcgacctg
ggctggcgcgacttcttcggcgaccgcgtgctgcaggagctgctggcgctgtcgctggcc
aacaaccgcgacctgcgggtcgccgcgcgtcaacgtggaggcgcgcgccctcaaccgcgagc
50 ggacaggccggcatcagccgcagctaccaggctcggtgccagcctgtcgacctgggagctg
gacctgttcggcgcatccgcagcctcagcgaacaggcgctgcagctctatctggcccag
gacgaaacgcgcctggccacccagctgacgctgggtggccgagaccgccaacgcctaccg
accctgcgcgcgcgaccaggaactgctggcgctgacgcgcgacgctggcgggccagcag
gagtcgtacaagctgaccgcgagactcagacctggcgctggcgaccgagctggacctg
agccggccgagatttcgtgcgcgaccgcgagcgcaatctgtcgagtagacgcgcgcatg
55 gcggcgcgaggaccgcaacgcgcgtggtgctgctgggtgggcccagccgctgccggccggcatc
ggcgcgagctggaccaggccgctggcgctgcccgcagggcgctggtcctggccgacctgccg
gcgggcctgcgcctcgatctgctgcgcgcgcccggcgatccgcgcggcgaggaccag
ctgcaagccgccaacgccagcatcggcgcggcgcgcgcgcggttcttcccgcgcacagc
ctgaccggctcgggcggcacggccagcgccagcctggcgggcctgttcgatgcggggtcg

5 ggggcctggagtttccgcccgcagatcagcgtgccgatcttcgcgggcgggcgctgcgc
 gccagcctggacctggccaagatccagaaggacatcggcatcgcgcgctacgagcaggcc
 atccagagcgggttccgcgaggtctccgacgcgctggcggcgcgccacattgcaggag
 cagatccgggtcgaggaactgctggtgcaggccaaccagcgccctacgacctgtcgag
 cagcgttaccagcagggcatcgacaactatctcagcgtgctggattcgagcgttcgctg
 tatacggcgagcagacgctggtcgagacggcgctggcgcgccctgtccaacctgatccag
 ctctacaaggcgctggcgcgggctggtccgagcgacgggtggcgggcgcgagggcggg
 tga

10 **SEQ ID NO:66 polypeptide sequence of Orf33**

MKPVVMRTLLSLAVATALAGCSLAPTYERPQAPVDAAYPSGPAYGAPGQAAAGAPAAADV
 GWRDFFGDPLLQELLALSLANNRDLRVAALNVEAARLNPSGQAGISRSYQVGASLSTWEL
 DLFGRIRSLSEALQLYLAQDETRLATQLTLVAETANAYPTLRADQELLALTRQTLAAQQ
 15 ESYKLTRQSYDLGVATELDLSQAEISLRTAERNLSQYTRMAAQDRNALVLLVGQPLPAGI
 GAQLDQAVLPDGVVLADLPAGLPDDLARRPDIRAAEHQLQAANASIGAARAAFFPRIS
 LTGSAGTASASLGLFDAGSGAWSFAPQISVPIFAGGALRASLDLAKIQKDIGIARYEQA
 IQSGFREVDALAGRGTLEQIERSQELLVQANQRAYDLSQORYQQGIDNYLSVLDSQRSL
 YTAQQTIVETRLARLSNLIQLYKALGGWSERTVAAAQAG

20 **SEQ ID NO:67 polynucleotide sequence of Orf34**

Atgaaacagcataaggtcggcaggcactgggcaggatgggcgatggcgctggcgctgcctg
 Ggcgcggcgccgcgcgctggcgcgagccggcgaccagctggggccgcgagggcgcg
 Gaactgctgctggaggtcaaggggcagcagcgttgcgcctggacgcgcgacctcgcg
 gtggcgatcgccgatccgcaggtcgccgacgtcaagggtgctggcgccccggcgctgggccc
 25 ccggcgaggtgctgctgatcgccggcaggccggcaccaccgagctgcgggtctggagc
 cgcggtcgcgcgacccgcaggtctggaccgtgcgcgtgctgccgaagtgcaggccgcg
 ctggcgcgcgcgcgctggcgggcgcgcgaggtcgacatggctggcgacagcgcgctg
 gtccaccggcatggcgccctcgcccgaggcgcatcgcgcgcgcgccgaggtgcccggggc
 gccgcgggcggaagtgaaagtgggtcgagctggcgcgctcggtcatgaaggatgtcgggatcaat
 30 ttccggggcgacagcgggcccggtggcgcgcgctcggtcatgaaggatgtcgggatcaat
 ttccggggcgacagcgggcccggtggcgcgcgctcggtcatgaaggatgtcgggatcaat
 ggccggcatgttcggcatgctgtcctataccagccgcgatttcagcgcgctcgctggcgctg
 ctgcaaaacaacggcatggcgcgcgctcctggccgagccgacgctgctggccatgtcgggc
 35 cagagcgccagcttctcgccggcgcgagattccgattccgggtatcgccggcgctgggt
 acgacctcggtgcagttcaagcccttcggcatcggcctgacgggtcacgcccacggtcatc
 tcgcgcgagcgcatcgcgctgaagggtggcgcccgaagccagcgagctggactacgccaac
 ggcatttccagcatcgacagcaacaatcgcatcacggtgatcccgcgcttgccaacccgc
 aaggccgacaccatggtggagctggcgcatggcgagacattcgctcatcagcgggcctgggt
 40 tcgcgcagaccaaaggccagcgtaacaagggtgcccgtggttggcgagacctgcccacatc
 gggcgcttcttcgcaacgtgcagttccagaggatcgcaattgggtgatcggtggtc
 acgcccgcgctggttcgccccatcgcgcgcggtgtcacgctgcccttgccggggcgcgcg
 caggaggtcagcgacgctggcttcaacgcctggggctattacctgctgggtccgatgagc
 ggccagcagatgccgggcttttcacagtga

45 **SEQ ID NO:68 polypeptide sequence of Orf34**

MKQHKVGRHWAGWAMALACLGAAPLAAQPAAPAGAAQARELLLEVKGQQPLRLDAAPSR
 VAIADPQVADVVKVLAPGVGRPGEVLLIGRQAGTTELRVWSRGRDPQVWTVRVLPQVQAA
 LARRGVGGGAQVDMAGDSGVVTGMAPSAEHRGAEEAAAAAGGNDKVVDMSQINTSGVV
 QVEVKVVELARSMKDVGINFRADSGPWSGGVSLLPDLASGGMFGMLSYTSRDFSASLAL
 50 LQNGMARVLAEPTELLAMSGQSASFLAGGEIPIPVSAAGLTTSVQFKPFGIGLTVTPTVI
 SRERIALKVAPEASELDYANGISSIDSNRITVIPALRTRKADTMVELGDGETFVISGLV
 SRQTKASVNVKPLGLDLP IIGAFFRNVQYSQEDRELVIIVTPRLVRPIARGVTPLPLGAR
 QEVSDAGFNWAGYYLLGPMMSGQQMPGFSQ

55 **SEQ ID NO:69 polynucleotide sequence of Orf35**

Atgaagcgacttctctgtctgtccctgctgtccgtattgctggcgggcgctgcacgacccca
 Tcgagattccgcccagacggcgcccgcgcgctgcgcggcgggcgccgaagggtccgctg
 Gtcgtgcccgcgctgtcgcgctgtccgacccccgcgcgcgctggcgggcgctac
 cagcgcggttgctggaccgagctgcccactgggagagcgacgacctgtcgcgctgggtgg

ccgctgttcctgcgcaattgcaaaggcctgatgcgccgaccagcggtaacctggcggcg
 ccggcacgcgccacgccgcgcctggcagcccgtgtgcgcggcgccgggtcgacccgtcc
 aaggcgccggcgccggcgacagcgccggcggtgcggcgcttccctgcagacctggctgcag
 ccctggcgcatcgccggcgccgacggcgctcccgccaccaataaccgtcaccggctactac
 5 gagccgctggtgcgcggctcgcgccgcccaggcgccgctaccagtggccgctgtatgcc
 gtgccggccgacctgctcgtcgtcgacctgggctcggtctatcccgacctgaccggcaag
 cgctgcgcggcgccgctcgacggccgcccgggtcgctgccctacgacacgcgcgcccgcgatc
 gagggggcgaccgcaagccgcccggccatcgctcgtgggtggacgatccggctcgacaatttc
 ttcttcgaggtccaggggtcgggcggggtgcagctgaccgatggccccgaccgcccacc
 10 acgatcccgctcggtacgcccagaccataacggccagccctatgcctccatcgggcgctgg
 ctcatcgacaagggcgagctgcgcgcccagaccggcatcgatgcagaacatccgtgcctgg
 gcccacgcaatccctcgcgctgcaggaaatgctcaacgccaacccggcggtggtcttc
 ttccgcaagaggcggtggtcgatccggagcaagggcccaagggggcctatggcatcccg
 ttggcgccgcagcgctcgatcgcggtcgacggcggttctgtgccgctgggacgcggctc
 15 tacctgtcgaccgctcgccgctccgacggccctcgacggccctgcagcgacccgtgttcgcgcag
 gacaccggcgccgacattcgcgccgcccgcgcgcccgaacttctattggggctacggcgag
 gaagccggccagcaggccggcgcatgaagcagcgccgagatgtggctgctgtggccc
 aagcaggccggggagccgtcggcgcgatga

SEQ ID NO:70 polypeptide sequence of Orf35

MKRLLCLSLLSVLLAACTTPSQIPPETAPGGVPPAAEGPLVVPPLSALSDTPPRALAGRY
 QRVAWTELPNWESDDLRSWWPLFLRNCKGLMRPTSGNLAAPARATPRAWQPVCAAVIDPS
 KAPAAGDSAAVRRFLQTLQWRIAGADGRPATNTVTGYEPLVRGSRROGGRYQWPLYA
 VPADLLVVDLGSVYPDLTGKRVGRDLGRRVVPYDTRAAIEAGDRKPPAIVWDDPVDNF
 25 FLQVQSGSRVQLTDGPDRTTIRVAYADHNGQPYASIGRWLIDKGELRADQASMQNIRAW
 AQRNPSRVQEMLNANPAVFFREEAVVDPEQGPKGAYGIPLAPQRSIAVDAGFVPLGTPV
 YLSTTLPASDRPLQRTVFAQDTGTAIRGAARADFYWGYGEEAGQQAGRMKQRGQMWWLLWP
 KQAGEPSAR

SEQ ID NO:71 polynucleotide sequence of Orf36

Atgttcaactgtcggcgattcctgcaaactcggcacgctgtcggccctgctggccggctgt
 Gccacctccagccaaacaccccaagcccagcatcttcccgcgccaggccgccacaggccag
 Gccgaccgctccgcacgcccggacaaacccgtatcgagcgacgaaggcccgccacg
 35 ctgacggcgaccggcgaaactgcggcccgacgtccgcgccttcgcccgaacagctggcgggc
 cagcgcgagctgcccctgcccgaagtgtggccagcctggaaagcacgcgctacaacgcg
 accgtcgcccgctcatcgccccgtccggcgcgctcgggcaagaaaatctggcgagctgg
 ctgacctatcgccggcggttctgtcgaaacccaagcgcatcgccctggggcggtggaattctgg
 aacgccaacaggacgtgctcaacggcgccgcccagcgctacggcggtgcccgcctcgatc
 atcgccctccatcatcgccgtggaaacccctgtatggccgcaacgtgggcaacttccgcgtg
 40 gtgcagcgcctggcgacgctggcattcgactacctcgatcccgccaagcccgagcgccg
 gacatgttcccgggccagctcgccgacttcatcaccctggcgctgcaggacaagctggac
 cccgagacgcgcggctcgtaacgcggcgccatcgccatgccgaattcatgcccggcagc
 atcatgcgctatcgccgtcgatggcgatgcgacggccacatcgacctgaccaacagcgctc
 gcgacgcggctcatgtcggtgggcaacttctgggtcgaaacatggctggcagcgccgctg
 45 ccgggtgttcgcgcgggtcgcgctgcccggccgatccggcgccgctgggtggccggcgccctt
 acgcccagcgctggactggaacggcctgcaggcccgccgcgcgcccgccggcgccgccc
 ggacggcgccgctggcaggagcaccatgggcacgtggacctggctgaggaagcgcgcc
 ggcaccgtgcaataaccgtaccgccacgccaatttctttgcccctgacgcaatacaaccgc
 agctacttctatgccacggcggtggccgacctggcgccgaactgcaggcccgccacgggc
 50 tattga

SEQ ID NO:72 polypeptide sequence of Orf36

MFNCRRFLQIGTSLALLAGCATSSQTPQAQHLPAQAATGQADRVIRIGPDKPVSSDEGPAT
 LTPTGELRPDVRAFAEQQLAAQRELPLPQVLASLESTRYNATVARLIAPSGASGKKIWRWSW
 55 LTYRGRFVEPKRIAWGVFEWVANQDLNRAAQRYGVPASIIASIIIGVETLYGRNVGNFRV
 VDALATLAFDYLDPAKPERADMFRGQLGDFITLALQDKLDPETRGSYAGAIGMPQFMPPGS
 IMRYAVDGGDDGHIDLTNSVADAVMSVGNFLVEHWQRGLPVFAPVALPADPAPLVAGGL

TPTLDWNGLAAGARPAAGAGRGAWQEHMPGIVDLVEEARGTVQYRTATPNFFALTQYNR
SYFYATAVADLAELQARTGY

SEQ ID NO:73 polynucleotide sequence of Orf37

5 Atgaaccatagactcatacgttgccctgagcatcgcgctgctggccctgctgtcgggctgc
Agcattctctccgggtcgggcccgcgacgcatcgcccatcatggacggcggtcgaccgac
Gcgaccggcgccaagctcggctcctacgacctgggtggacctgcgcgcccacaccattgcy
ccctatgtgctggtcaaggcgggtgtccaaggatggcgccacctcgacgggtacgtgggc
aatatgcgcgtgatgccgggcatgtgctgcgcacctggtagccgacagcatggagacc
10 ggactgttcgcgcgctggcggcgccggcgccgacgggtgttcgaagccgtcggggtcgcgcc
gacggcagcatctcgctgccctatgcgggcccgcctgaaagtgcagggcaagtgcgtggcg
cagatcgagcagctcgtcaagggcagcctgcgcaataccgcggcggtgcagccgcaggcc
atgggtggatctggccgacgaccgctccaattcggtgctggtggccggggcggtgccgcgc
ccgggacgcttcggcgcccaacaaggcccgcgtgacggcgctggatgcgatcacgcaggcg
15 ggcggctcgaccctgccggcttaccaggccgacgtagtgcggactggcagcaagggtg
cagcgcatctccttaccagcaattgctcaacggccgcaacgtggcggtggagccgcgctcc
gaactggtggtcgaaccgaacctgaagcgttctcgtggcgatgggggcccctaccaagccg
ggcctgcacgaactgccgtcgaaaccagaccaatctgctcgacgcccctggcggtggccgga
ggcctgaacgacgcgcggcgccgacgccaccggggtattcggttttcgcctggacggccgc
20 aacgccgatggccgcccgcggccacgggtgttcaggctgaatatgcgcaatccggagtcc
atgttcctggccaagcaattcgagctgctgcggaggacgtggtgtatgtcagtaatgcg
cccatgtacgaatgggaaaagatcattacgcctatcgtgcaggctcctgatcgtgggcca
cgcgtgggtacttactaa

25 SEQ ID NO:74 polypeptide sequence of Orf37

MNHLIRCLSIALLALLSGCSILSGSGPTRSAIMDGGSTDATGAKLGSYDLVLDLRADTIA
PYVLKAVSKDGATSDGYVGNMRVMPGDVLRILVADSMETGLFAPLAAGGTVFEAVRVAA
DGSISLPYAGRLKVQKSLAQIEQLVKGSLRNAAVQFPQAMVDLADDRSNSVLVAGAVPR
30 PGRFGGNKGPLTALDAITQAGGSTLPAYQADVIRTGSKVQRIPYQQLNGRNVAVEPRS
ELVVEPNLKRFFVAMGALTKEPLHELPSNQTNLILDALGVAGGLNDRAADATGVFVRLDGR
NADGRPRPTVFRLNMRNPESMFLAKQFELLPEDEVVYVSNAPMYEWEKIITPIVQVLIVGQ
RVGTY

SEQ ID NO:75 polynucleotide sequence of Orf38

35 Atgaacgtctcatgcccacacctgggtcggaactgctcgctcctggccgctcctgtcttca
Tgcgtcttcgtgggtccgcgagcgcgactacgccctgggtgttctcgctggggcgaggtgcgc
Caggtcatcagcgagcctggcctgtatttcaaggcgccgcgcgcttcagaaagctcgctc
acgctggacaagcgcatcctcaccatcgagtcagcgatgccgagcgcatccagacctcc
40 gagaagaagaacctgctgatcgactcgtaactgagtcgaagtcgagcgcgcatccgcgcctg
tactacgtgaccttcggcggaacgagcgcgccgcccaggagcgtctgcaggcgagatc
cgcgacgcgctgaacgcggcggtcaacgtgcgcacgggtcaaggacgtggtctcgcccgag
cgtgacaaggctcatggccgaaatcctcaccaacgtcgtcaagcgcgccgagcgctgggc
gtgcagggtggtcgacgtgcgcctgcgcgcgcatcgagttcgcgcccagagatttcgagtcg
50 gtctatcgccgcatggaagccgagcgacccgcgtggccaacgagctgcgttcgatcggc
gcccgcgaaagcgagaagatccgcgcgagggccgacggccagcgcgaggtcatcgtggcc
caggcctatgcgcgcgcccaggggcatcatgggagggggcgacggccaggccggcagcatc
tacggccaggccttcggcgcaataccgagttctacacctattacaagagcctggaagcc
tatcgcgccgcttcggcaaaaccgggtgacgtattgggtggtcgatccgacgtcggagttc
60 ttccagttcttcaagaaccccggaaggcgcgggcgcccgccgacccggcgaattga

SEQ ID NO:76 polypeptide sequence of Orf38

55 MQRLMPILVGLLVLAVALSSCVFVVRERDYALVFSLGEVRQVISEPGLYFKAPPPFQNVV
TLDKRILTISSDAERIQTSEKKNLLIDSYVKWRIADPRLYYVTFGGNERAAQERLQAI
RDALNAAVNVRTVKDVVSAERDKVMAEILTNNVVKRAEPLGVQVVDVRLRIEFAPESIS
VYRMEAEERTRVANELRSIGAAESEKIRAEADRQREVIVAQAYARAQIMGEGDAQAGSI
YAQAFGRNTEFYTTYKSLEAYRAAFGKTGDVLVDPTSEFFQFFKNPGKAAGAPAPAN

SEQ ID NO:77 polynucleotide sequence of Orf39

Ttgcccagggaggcaaccatgaaacccgtcatccagactttcctgcgcgcgcgcgcgcgtg
 Gccggcctggcgctgctggccggctgcgcgcgcgtcagcacgacgcagtcgcgcgcgcgatc
 Ggcgtggaccgcacccaatacatgtcgagcctgggtgcccgagcaggcgcgtggtgcaggag
 gccgggcagcagtatgccgagatcgtccaggaggcccgcgccaaggggctgcttgaccgc
 5 gacccggcgcaattgtcgcgctgcgcgccatttcccagcgctgatcgcgagaccggg
 gtgtttcgcgccgacgcggccaactggccatgggaagtgcagtgctgctgcgtcgacgag
 gtcaacgcctgggtgcatgccggcggaagattgccgtctacacgggcctgctcgccat
 atcaagccgacgcagcgaactggcgcggtgctgggcccagagatcgcgcatgctgctg
 10 cgcgagcacgcgcgcgcgcgtctcgagcagatggcgaccagcatcggcctgctcggtg
 ctgtccatggccaccggttgcgcgcgcgtccgacctggcgcggaagctgaccgaagtc
 atgttcaccttgcaccaacagccgcacgcacgagaccgaggccgatcgcatggcgctcgaa
 ctggcgcgcgcgcgcggttccgatccgcgcgcgcgcgtcacgctgtggcagaaaatgggc
 gggcgacggcaatgcgcgcgcggagttcctgtccacccacccgtcgccagtagccgc
 15 atcggcgaattgcagcaggccttgcaagaaggtattgccgctgtacgagcaggcgcgcgcc
 caggccgccaatag

SEQ ID NO:78 polypeptide sequence of Orf39

LPREATMKPVIQTFLRAAAVAGLALLAGCAGVSTTQSGAIGVDRTQYMSSLVPEQALVQE
 AGQQYAEIVQEARAKGLLDRDPAQLSRVRAISQRLIAQTGVFRADAANWPWEVHVLSDVE
 20 VNAWCMPPGGKIAVYTGLLAHIKPTDDELAAVLGHEIAHALREHARERSVQOMATSIGLSV
 LSMATGSPGASDLGGKLTVMFTLPNSRTHETeadrmgvelaaragfdpRAAVTLWQKMG
 AADGNAPPEFLSTHPSASTRIGELQQALQKVLPLYEQARGQAAK

SEQ ID NO:79 polynucleotide sequence of Orf40

Gtgactcaccgtcccgtgcactctcgaagcccgcctcccgcgcgcggggtggccctgcgc
 Gcggcgatcgcgctgtcaaccattctgatcgtggccggctgcggctcgtaagcaccaaa
 Tacgacaagaccgcgggctggagcgccgaacagttgtacgccgacgccaagcaggaagt
 ggcgggggcaactggaccgatgcccgggagcgccctgaccgcatcgaaagccgctaccg
 25 ttcggcacgtacgcccagcaggccctgatcgaactggcttacgtcaactggaaagacggc
 30 gagaacgaacaggcgctggcgccatcgaccgcttccagcagctctatcccaaccacccg
 ggcacggactacgtgctgtacctgaaggggctgggtcaacttcacgcccggccagcgcttc
 atgagcaacctgaccggccaggaccccgcgcgagcgatcccaagggcctgcgcgcgtcc
 tacgatgcgttcaacgaactgggtccagcgcttccccaacagcaagtacacgcccgatg
 35 cagaagcgcatgacctggctgggtcaacgcatcgccatgaacgaagtccacgtggcgcg
 tactactacgagcggggcgccctacgtggcgccgccaaccggggcgagaccgtgatcacc
 gatttcgagggggggggcgccctcggaagaagcgctctatatcatggtcgagtcgatgac
 aagctgggaatgaccgaactgaagggcgacgcccgaacgcgtgctcgaccagaactatccc
 aacagcaaattcaagacgaaggcctgtcgggccgacaagagctgggtggaaccgcttctcg
 40 tggcgctga

SEQ ID NO:80 polypeptide sequence of Orf40

VTHRPAALSKPASRRGVALRAAIALSTILIVAGCGSSSTKYDKTAGWSAEQLYADAKQEV
 AAGNWTDAERERLTAIESRYPFGTYAQQALIELAYVNWKDGENEQALAAIDRFQQLYPNHP
 45 GTDYVLVYLKGLVNFTPASAFMSNLTGQDPAERDPKGLRASYDAFNELVQRFPSKYTPDA
 QKRMTWLVAIAMNEVHVARYYYERGAYVAAANRAQTVITDFEGAPASEEALYIMVESYD
 KLGMTTELKGDARVLDQNPNSKFKTQGLSADKSWWNPFSWR

SEQ ID NO:81 polynucleotide sequence of Orf41

Ttgccccacaggttgaccttgccatgacgaagcactctgccgctcgaatcgccaccatc
 50 Gccgcgcgagggcgtcctgctggccggctgcgcagcgcccaagaaccccgatccgcgcgat
 Ccctgggaaggcttcaaccggggcgtctacaagttcaacgacacggctcgaccgcgcgctg
 ttaagccgggtggcccaggcctataccttcgtcaccccgagcgggtgcgcagctgcgtg
 cacaatatgttcagcaacgtggcgacctgtggctcgccaccaacagcttccctgcaaggc
 cgcgggcacgatttcgtcaacacgatcgccgcgttccctgttcaataccaccatgggggatc
 55 ggcggtgcttcgacgtcgcgtagaccggggcgcgcaagatcccaacgacttcggc
 gtgacgtggcgctctggggcttcggccagggacgtacctgggtgctgccgatctggggc
 gccagcagcctgcgcgacggcgctcgccctgatcggcgactggaccgggcaaccaggggcg
 accatcggcgcgatcgacaacgtgccgctgcgcaactcgctgtggggcctggaggccgtc

gacctgcgcgccagcctgctcgataccaccgacaccgtggaccgcgtggcgctggatccc
 tacagcttcgtgcgcgacgcctacctgcagcgccgcgcgccatggtgcgcggcaccacg
 acggcgacgacacgctgcccacctatgaagacgagggcgatgacgacgcgcccccgcc
 gcgcggccgcccagcggccgcccagcgcagtaa

5

SEQ ID NO:82 polypeptide sequence of Orf41

LPPQVDLAMTKHSAARIATIAAAGVLLAGCAAPKNPDPRDPWEGFNRGVYKFNDTVDRAL
 FKPVAQAYTFVTPQPVRSVHNMFSNVGDLWSATNSFLQGRGHDFVNTIGRFLFNTTNGI
 GGCDFVASTTGARKIPNDFGVTLGVWFGQGPYLVLPWIGASSLRDGVGLIGDWTGNQGA
 TTGAIDNVPLRNSLWGLEAVDLRASLLDTTDTVDRVALDPYSFVRDAYLQRRAMVRGK
 TGDDTLPTYEDEGDDDAAPAAPAAQPAAPQ

10

SEQ ID NO:83 polynucleotide sequence of Orf42

Atggcaacaaagtgcctgctccaggggagttttccggatgccagcccgataatgccggca
 Atgcgtagtggcgccgcatgggtgctggaaggaggtttatgcggttggatggggattg
 Ccggcgctggccgtcgtgcttgcgctggccggatgcgtgaatcgcgagccagaggagcgc
 gcggccttcacgcgtatctggaacaagtggccgcgcgcgagggcggtcgtggccgcg
 ccgcccgcacccgcccacgcgcaaggccctggcgactacgaggcgagtagcagccgatg
 gaagcggcgacgcgcgcgtgcgcgaagcgttggcgccgcagcagggcggtcgcaggcg
 ctgcggctgcattcggtcgacgagatcgtcgacgcgcagcagggctgggacaggctggcc
 gagcgctggcgccgcgcgcacccgggtcgaacaggcgcgccgcgcgcgcgcgcgcgcgc
 cgcgccgggatggagcagcctcccagcctgcgcaacgcctacgcgcgcgcctatgaacac
 agcgtcacggcgccggcacaggccttggcgcgatataccggcctgctcgaaccgcgcgtg
 gaggatgcgcggcggtggccgggttcggtgcgcgccatcgcgatcaggtcgataccgat
 ggtccgctgaccaggtgcgcgatccctcggtgcgcagcgagctcaatgtactgctgcag
 gcgctcaatggcgctccgaccaggttcgcaggcgagccttgcctcaatggcctggcg
 ggaccggctcgccaggcgccctga

15

20

25

SEQ ID NO:84 polypeptide sequence of Orf42

MATKCLLQGSFPDASPIPMRSGAAVWLEGRFMRFGWGLPALAVVLALAGCVNREPEER
 AAFIAYLEQVAAPQAGVVAAPPDPPTRKALGDYEAQYEPMEAAHAAVREALAAQQAALQA
 LRLHSVDEIVARQDQWDRLEAERLAAARTGLEQARAAADAARAGMEQPPDLRNAYARAYEH
 SVTAPAQALARISGLLEPAVEDARRVAGFVARHRDQVDTDGPLTQVRDPSVRSELNVLLQ
 ALNGRSDQVSQAQALLNGLAGPARQAP

30

35

SEQ ID NO:85 polynucleotide sequence of Orf43

Gtgatgctgaagaccgtattgcgcctgcgggtcgtgcgcgcgctgctggcgctggccgcg
 Ggctgcgcgatgattccgcgccgaaccgggtggtgatctgtccgctgaccgcgcgcgcctccg
 Tgcgcgcgcgcaaccctcgccgcggcccaacggctcgatctaccagccttcggcctacggc
 aactatccgctgttcgaggaccgcgcggccgcgcgaacgtggcgacatcgtcaccatcggtg
 ctggaggaaaagaccaacgcgcgaagggcggtggccaccaataccagccgcgcgcgcgcgc
 gccacgctggcggtggcgccgcgcgcgcgcgcgttcattggacggcatcatcaacgacaagctg
 gataccgatatactcgggcggaataaccgccaacggcaccggcaagagcagcgccaacaac
 accttcaccggcaccatcacgaccaccgtgatcggggtgctgccaacggcaatctgcag
 atcgccggcgagaagcagatcgccatcaaccgcggcagcgagtagctgcgcttctcgggc
 gtggtcgaccgcgcgatcgatcacggcgagcaatacgggtgctcgaccgggtggccgac
 gcgcgcgatcgaataccgcagcaagggcgctcatggacgaagtcagaccatgggctggctg
 caacgctttttcctgatcgcttcgccgttctga

40

45

SEQ ID NO:86 polypeptide sequence of Orf43

VMLKTVLRLPVCAALLALAAGCAMIPPEPVVICPLTAPPPSPQPSARPNGSIYQPSAYG
 NYPLFEDRRPRNVGDIVTIVLEEKTNAAKGVAATNTSRDGSATLGVAAPRFMDGIINDKL
 DTDISGGNTANGTGSSANNTFTGTITTTVIGVLPNGNLQIAGEKQIAINRGSEYVRFSG
 VVDPRSITGSNTVSSTRVADARIEYRSKGVMDDEVQTMGWLQRFLLIASPF

50

55

SEQ ID NO:87 polynucleotide sequence of Orf44

Atgaagtgcctccctgtatcgaatcgacgcgctcagcgccgctgcctgttgctggccggc
 Tgcgcccaaccagcgcgctccgaaggagtcggggttcctcggcgattactcgagttgcgc

Gaggagcaggtgccccggcgcgcgcggtgatctaccgcgacgcccgcgctcaagccgcg
 cagtacaccgccatgtggctgtcgccggtcgagtactacccagcccgcaaccgtcggcg
 caggtgtcgatggaaacgctgaccgaactgcagaactacctggaccagtcgctgcgccgc
 aagatcgcccgcgagatccgcctggtcaacggccccggccccggcggtggccaaggcgcg
 5 atcgcgatcacagcggtcggcagcgaaagcgaggcgctggcgccctaccagtacatcccc
 gtggcgctggccgtcaccggcgccagggccgtgctggaaggcgccggcgcgagcaggcc
 accatcgcgatcgaaagcaaggtcaccgacagccagacggggccagctgctgtggcgctcg
 gtgcgccccggcgagcgcgtagcgccatcgcccagggccaggcctcggtgccc
 10 gcctcgcgctcaagccgctgatcgacgaatggaccgataacgtcgcacgtgaaatacgc
 aactacgtgcgagcaataa

SEQ ID NO:88 polypeptide sequence of Orf44

MKSSLYRIAALSAAALLLAGCANQRAPKESGFLGDYSQLREEQVPGGARLIYRDAALKPR
 QYTAMWLSFVEYYPSQPQSAQVSMETLTQLNYLDQSLRRKIGREIRLVNPGPGVAKAR
 15 IAITAVGSESEALAAYQYIPVALAVTGARAVLEGGRPQQATIAIESKVTDSTGQLLWAS
 VRGGTGERVRAIAQQQASVPASALKPLIDEWTDNVAREIRNYVRSK

SEQ ID NO:89 polynucleotide sequence of Orf45

Gtgaaccaacgtggggcccttttaccggttaacacgtgtgactctctttgcaaaggaact
 20 Atcatgaagtcgcgcatcgccaaaagcctaaccatagctgcgctggccgccacgtggca
 Gcctgcagttccgtccctctcgacgacaaggcaggtcaagctggaggctccggccagggt
 tcggcctccggccagatcctggatcccttcaaccgcaaaagcattctggcgcaacagcgc
 tcggtgtactttgacttcgacagctatacgggtgctcggaacagtatcgcgccctggtcgaa
 25 acccacgccccgtacctggcttcgaacaaccagcagcgcatcaagatcgaaggcaatacc
 gacgaacgccccggcgccgagtagaacctcgactgggccaacgccgtgccgacgtgtc
 cgctcgcatgatgacctgctgggtgtgtcggaacaaccagatcgaaaccattagtttcggc
 aagggaaagccgaaggcgacgggttcgagcgaggctgatttcgccgagaaccgcccgcgcc
 gatatcgtttatcagcgctaa

SEQ ID NO:90 polypeptide sequence of Orf45

VNQRGALLPVNTCDLCKGTIMKSRIAKSLTIAALAATLAACSSVPLDDKAGQAGGSGQG
 SASGQILDPFNPQSILAQQRSVYFDSDSYTVSEQYRGLVETHARYLASNNQORIKIEGNT
 30 DERGGAEYNLALQRRADAVRRMMTLLGVSDNQIETISFGKEKPKATGSSEADFAENRRA
 DIVYQR

SEQ ID NO:91 polynucleotide sequence of Orf46

Gtgtccatgatcgacgtatttccctcgccgctctgaaggggctcgcggtggctgtcctg
 Gcagcctccgccctgaccgctgctcgctcgccgcaaatggggattcccctacaaggccggc
 40 Gtccagcaaggcaactggatcacaaagagcaggtcgccctgctgcagcaaggcatgtcg
 cgcgaaacaggtgcgcttcgccctgggcagccccacgtgaccagcgtgctgcacgccgat
 cgctgggattaccctactacttcaagcccggtacggcaaggcgaggaacgccagttc
 accgtgtggttcgagaacgaccacctggtagctggagcggggatgaacagccccgacctc
 cagccgttccagatcgagaaagtgaacgccaacaggaagaaaaagccgacgcccaggtg
 45 gatacggcccgagaagcgccaggaaggcatcgacaaggctgaaaaagtccggccccatgtc
 gatgtcacgacgcccggacaacccccacctcgactaccgggagcgccgggccaacacctc
 gaaccgctcaagtaa

SEQ ID NO:92 polypeptide sequence of Orf46

VSMIARISLRPLKGLAVAVLAASALTACSSGKWGFYKAGVQQGNWITKEQVALLQQGMS
 50 REQVRFALGSPTLTSLVHADRWDPYFYFKPGYGKAQERQFTVWFENDHLVRWSGDEQPD
 LQPFQIEKVNAKQEEKADAQVDTAEKRQEGIDKAEKVRPHVDVTTDPNPTLDYPGEPGQTF
 EPLK

SEQ ID NO:93 polynucleotide sequence of Orf47

Atggcgaccatcctgtcgggccaacgttgctggcgcgctgacgctgcttgccgcctgc
 Agcggttccatggcgcaagagccgcctacaagagcacgatactgggcttgaggcgacc
 55 Atcctggacctgaagggttgccgtccgcacacggcgcgcatatcgacacctgagcgcc
 caagtgggtgctgctggccgcgcccatgaaggcggtgctcggtacggcgagggaaggatgcc

gtcaccatcgccatgatggcgacgtactcttcgatttcgacaaggccgacatactcgcc
 gcggccgaacccactctgcgggacatcgcgagctgatcaaatacccccgccaccggcatc
 gtgcgcattgaaggtcacacggactccaagggtcggattcctataacaagggcctgtca
 ttgcgacgggcccaggccgttgcgcagtggtggcgctcacggggtggatgcagcgaaa
 5 ctgtcggtcaggggcctgggggctgccaggcccgtacagcccaaccagctagctgtgaag
 attcaatag

SEQ ID NO:94 polypeptide sequence of Orf47

MATHPVGPPTLLAALTLAACSGSMAQEPYPKSTILGLQATILDLKGLPSDTDGGISDL
 10 QVGALAARHEGVSVRQKDAVTIAMMGDVLFDKADILAAEPTLRDIAELIKSPATGI
 VAIEGHTDSKGSDSYNKGLSLRRAQAVAQWLGAHGVDAAKLSVRGLGAARPVQPNQLAVK
 IQ

SEQ ID NO:95 polynucleotide sequence of Orf48

Atgaactatatgcattccccctctgtagttgccgggcgcgcccgccgcctgctggcggt
 15 Gcgggcggttgccgggtcgggtggccgttctggccggctgcgccaatcccagcgcatcgagt
 Ggggtgtacacgtacggccaggcgcagcgcgagcagatcgtgcgcaccgggcacggtcacc
 ggcgtgctccgattaccatccagaacgacaagtccagcggtcgtcggttgggtggccggt
 20 ggcgcgctggggcggttagcgggcaatgccgtcggcgggcggcaccggccgcaccatcgcc
 acgggtggggcggtcatcctcggcgcgctggcgggcaacgccatcgagaaccgcgcgggc
 aagtcctccggctacgaaatcacgggtgcgcctggacaacggcgaaacccgggtcgtggcg
 caggaagccgacgtgcccatcagcgtggggcagcgcgtgcagggtcatcagcggcgcgggc
 ccgacccgcgtgacaccgtattga

SEQ ID NO:96 polypeptide sequence of Orf48

MNYMHSPSVVAGRARRLLAVAAVAGSVAVLAGCANPSASSGVYTYGQAQREQIVRTGTVT
 25 GVRPITIQNDKSSGVGLVAGGALGGVAGNAVGGGTGRTIATVGGVILGALAGNAIENRAG
 KSSGYEITVRLDNGETRVVAQEADVPIISVGQRVQVISGAGPTRVTPY

SEQ ID NO:97 polynucleotide sequence of Orf49

Ttggcggtgatcagcaaaaaggagcgcacatcttgaaaaaccctgctacccgtattggcgctt
 30 Gccgccctgctgtcggcctgcaacgcgaacgccccctcggatacgcggaggggcgcgccg
 Ccgcccgatacgcatacctcgcgcaattcgctggactggcaaggcacgtaccagggcggtg
 ctgccgtgcgcgactgccccggcatccgcacgggtgctgaccctgcgcgcccgaacaacc
 35 taccagttgcagaccagtagctggagcgccagccccgcccggacacgggtgcaaggcaga
 ttccggctgggtgacggggcgaacgccatcgagctcgacagcgccggcgatcactaccgt
 taccaggtcggcgaaaaccgggtgaccatgatgtcgcaagacggcacccctgccagcggc
 ccgttggccgagcactacgtgctcaagcgcagccagtga

SEQ ID NO:98 polypeptide sequence of Orf49

LALISKKERILKTLPLVLALAAALLSACNANAPSDTPEGAPPPDTHTSRNSLDWQGT
 40 LPCADCPGIRTVLTLRADNTYQLQTQYLERQPRPDTVQGRFGWLTGDNAIELDSAGDHYR
 YQVGENRLTMMSQDGTLPSPGLAEHYVLKRSQ

CLAIMS:

1. An immunogenic composition comprising a polypeptide comprising an amino acid sequence which has at least 85% identity to an amino acid sequence selected from the group consisting of SEQ Group 2 , over the entire length of said sequence from SEQ Group 2, or an immunogenic fragment thereof and a pharmaceutically acceptable excipient .
2. The immunogenic composition as claimed in claim 1 in which the polypeptide comprises an amino acid sequence which has at least 95% identity to an amino acid sequence selected from the group consisting of SEQ Group 2, over the entire length of said sequence from SEQ Group 2, or an immunogenic fragment thereof.
3. The immunogenic composition as claimed in claim 1 in which the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ Group 2, or an immunogenic fragment thereof.
4. The immunogenic composition of claims 1-3 in which the polypeptide comprises an amino acid sequence which is an immunogenic fragment of the polypeptide of SEQ Group 2 in which the immunogenic activity of said immunogenic fragment is substantially the same as the polypeptide of SEQ Group 2 .
5. The immunogenic composition as claimed in any of claims 1 to 4 wherein said polypeptide is part of a larger fusion protein.
6. An immunogenic composition comprising a polynucleotide comprising a nucleotide sequence having at least 85% identity to the nucleotide sequence of SEQ Group 1, over the entire length of said sequence from SEQ Group 1 and a pharmaceutically acceptable excipient.

7. The immunogenic composition of claim 6 wherein the polynucleotide comprises a nucleotide sequence having at least 95% identity to the nucleotide sequence of SEQ Group 1, over the entire length of said sequence from SEQ Group 1.
- 5 8. The immunogenic composition of claim 6 or 7 wherein the polynucleotide comprises a nucleotide sequence having the sequence of any polynucleotide selected from the group consisting of SEQ Group 1.
- 10 9. An immunogenic composition comprising a polynucleotide encoding a polypeptide comprising an amino acid sequence which has at least 85% identity to an amino acid sequence selected from the group consisting of SEQ Group 2, over the entire length of said sequence from SEQ Group 2 or an immunogenic fragment thereof, and a pharmaceutically acceptable excipient.
- 15 10. The immunogenic composition of claim 9 wherein the polynucleotide encodes a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ Group 2, or an immunogenic fragment thereof.
- 20 11. An immunogenic composition comprising at least or exactly two, three, four, five, six, seven, eight, nine or ten different Bordetella, preferably *B. pertussis* antigens wherein the antigens are selected from at least two, three, four or five groups of proteins selected from the following:
 - a) at least one Bordetella iron acquisition protein selected from the group consisting of
25 a polypeptide sharing at least 70%, 80%, 90%, 95%, 97%, 98%, 99% or 100% identity with SEQ ID 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, or 28, or an antigenic fragment thereof;
 - b) at least one Bordetella autotransporter protein selected from the group consisting of pertactin, BipA, a polypeptide sharing at least 70%, 80%, 90%, 95%, 97%, 98%,

99% or 100% identity with SEQ ID 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, or 54, or an antigenic fragment thereof, preferably the passenger domain thereof;

- c) at least one Bordetella lipoprotein selected from the group consisting of a polypeptide sharing at least 70%, 80%, 90%, 95%, 97%, 98%, 99% or 100% identity with SEQ ID 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, or 98 or an antigenic fragment thereof;
- d) at least one Bordetella adhesin selected from the group consisting of FHA, fimbriae 2 and/or 3, pertactin and BrkA or an antigenic fragment thereof; and
- e) at least one Bordetella toxin/invasin or antigens involved in toxin/invasin secretion selected from the group consisting of pertussis toxin, adenylate cyclase, dermonecrotic toxin (Dnt), Type III ss or lipopolysaccharide or an antigenic fragment thereof,

wherein the Bordetella antigens in the immunogenic composition do not consist of any combination of 2, 3, 4 or all 5 of pertactin, fibriae 2, fimbrae 3, FHA and pertussis toxin.

12. The immunogenic composition of claim 11 comprising one or more Bordetella iron acquisition protein selected from the group consisting of the polypeptides sharing at least 70%, 80%, 90%, 95%, 97%, 98%, 99% or 100% identity with SEQ ID 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26 or 28, or an antigenic fragment thereof.

13. The immunogenic composition of claim 12 wherein the Bordetella iron acquisition protein is BhuR or an antigenic fragment thereof.

14. The immunogenic composition of claim 11-13 comprising a Bordetella autotransporter protein selected from the group consisting of pertactin, BipA and a polypeptide sharing at least 70%, 80%, 90%, 95%, 97%, 98%, 99% identity with SEQ ID 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, or 54, or an antigenic fragment, preferably the passenger domain, thereof.

15. The immunogenic composition of claim 14 wherein the Bordetella autotransporter protein is BipA, or an antigenic fragment, preferably the passenger domain thereof.
- 5 16. The immunogenic composition of claim 14 or 15 wherein the Bordetella autotransporter protein is pertactin, or an antigenic fragment, preferably the passenger domain thereof.
- 10 17. The immunogenic composition of claim 14-16 wherein the Bordetella autotransporter protein is BrkA, or an antigenic fragment, preferably the passenger domain thereof.
- 15 18. The immunogenic composition of claim 14-17 wherein the Bordetella autotransporter protein is TcfA, or an antigenic fragment, preferably the passenger domain thereof.
19. The immunogenic composition of claim 14-18 wherein the Bordetella autotransporter protein is BapA, or an antigenic fragment, preferably the passenger domain thereof.
- 20 20. The immunogenic composition of claim 14-19 wherein the Bordetella autotransporter protein is BapB, or an antigenic fragment, preferably the passenger domain thereof.
- 25 21. The immunogenic composition of claim 14-20 wherein the Bordetella autotransporter protein is BapC, or an antigenic fragment, preferably the passenger domain thereof.
- 30 22. The immunogenic composition of claim 14-21 wherein the Bordetella autotransporter protein is YapE, or an antigenic fragment, preferably the passenger domain thereof.

23. The immunogenic composition of claims 11-22 comprising a Bordetella lipoprotein selected from the group consisting of BipA, the polypeptide sharing at least 70%, 80%, 90%, 95%, 97%, 98%, 99% identity with SEQ ID 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96 or 98 or an antigenic fragment thereof.
24. The immunogenic composition of claim 23 wherein the Bordetella lipoprotein is MltA or an antigenic fragment thereof.
25. The immunogenic composition of claim 23-24 wherein the Bordetella lipoprotein is MltB or an antigenic fragment thereof.
26. The immunogenic composition of claim 23-25 wherein the Bordetella lipoprotein is VacJ or an antigenic fragment thereof.
27. The immunogenic composition of claim 23-26 wherein the Bordetella lipoprotein is OmlA or an antigenic fragment thereof.
28. The immunogenic composition of claim 23-27 wherein the Bordetella lipoprotein is Pcp or an antigenic fragment thereof.
29. The immunogenic composition of claims 11-28 comprising a Bordetella adhesin selected from the group consisting of FHA, fimbriae, pertactin and BrkA or an antigenic fragment thereof.
30. The immunogenic composition of claim 29 wherein the Bordetella adhesin is FHA or an antigenic fragment thereof.

31. The immunogenic composition of claim 29-30 wherein the Bordetella adhesin is Fimbriae 2 and/or 3 or an antigenic fragment thereof.

32. The immunogenic composition of claims 11-31 comprising a Bordetella toxin/invasin or antigens involved in toxin/invasin secretion selected from the group consisting of pertussis toxin, adenylate cyclase, dermonecrotic toxin (Dnt), Type III ss or lipopolysaccharide or an antigenic fragment thereof.

33. The immunogenic composition of claim 32 wherein the Bordetella toxin/invasin or antigen involved in toxin/invasin secretion is pertussis toxin or an antigenic fragment thereof.

34. The immunogenic composition of claim 32-33 wherein the Bordetella toxin/invasin or antigen involved in toxin/invasin secretion is Type III ss or an antigenic fragment thereof.

35. The immunogenic composition of claim 32-34 wherein the Bordetella toxin/invasin or antigen involved in toxin/invasin secretion is LPS or an antigenic fragment thereof.

36. The immunogenic composition of claim 11-35 comprising a) FHA, b) pertussis toxin and c) BrkA or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 34, or an antigenic fragment thereof, preferably further comprising d) pertactin.

37. The immunogenic composition of claim 11-36 comprising a) FHA, b) pertussis toxin and c) BhuR or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 20, or an antigenic fragment thereof, preferably further comprising d) pertactin.

38. The immunogenic composition of claim 11-37 comprising a) FHA, b) pertussis toxin and c) BapA or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 40, or an antigenic fragment thereof, preferably further comprising d) pertactin.

5

39. The immunogenic composition of claim 11-38 comprising a) FHA, b) pertussis toxin and c) BapB or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 42, or an antigenic fragment thereof, preferably further comprising d) pertactin.

10

40. The immunogenic composition of claim 11-39 comprising a) FHA, b) pertussis toxin and c) BapC or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 44, or an antigenic fragment thereof, preferably further comprising d) pertactin.

15

41. The immunogenic composition of claim 11-40 comprising a) FHA, b) pertussis toxin and c) YapE or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 52, or an antigenic fragment thereof, preferably further comprising d) pertactin.

20

42. The immunogenic composition of claim 11-41 comprising a) FHA, b) pertussis toxin and c) VacJ or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 82, or an antigenic fragment thereof, preferably further comprising d) pertactin.

25

43. The immunogenic composition of claim 11-42 comprising a) FHA, b) pertussis toxin and c) OmlA or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 92, or an antigenic fragment thereof, preferably further comprising d) pertactin.

30

44. The immunogenic composition of claim 11-43 comprising a) FHA, b) pertussis toxin and c) Pcp or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 96, or an antigenic fragment thereof, preferably further comprising d) pertactin.

5

45. The immunogenic composition of claim 11-44 comprising a) FHA, b) pertussis toxin and c) MltA or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 70, or an antigenic fragment thereof, preferably further comprising d) pertactin.

10

46. The immunogenic composition of claim 11-46 comprising a) FHA, b) pertussis toxin and c) MltB or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 72, or an antigenic fragment thereof, preferably further comprising d) pertactin.

15

47. The immunogenic composition of claim 11-46 comprising a) FHA, b) pertussis toxin and c) TcfA or a protein sharing at least 70, 80, 90, 95, 97, 98, 99 or 100% identity with SEQ ID 36, or an antigenic fragment thereof, preferably further comprising d) pertactin.

20

48. The immunogenic composition of claim 11-47 comprising a) FHA, b) pertussis toxin and c) adenylate cyclase, or an antigenic fragment thereof, preferably further comprising d) pertactin.

25 49. The immunogenic composition of claim 11-48 comprising a) FHA, b) pertussis toxin and c) Type III ss or an antigenic fragment thereof, preferably further comprising d) pertactin.

50. The immunogenic composition of claims 1-49 for the prevention or treatment of *Bordetella pertussis* infection.

30

51. The immunogenic composition of claims 1-50 for the prevention or treatment of *Bordetella parapertussis* infection.

5 52. The immunogenic composition of claims 1-51 for the prevention or treatment of *Bordetella bronchiseptica* infection.

53. The immunogenic composition of claims 1-52 comprising a polypeptide that is expressed during the Bvg⁺ early phase of *Bordetella* infection.

10

54. The immunogenic composition of claims 1-53 comprising a polypeptide that is expressed during the Bvg⁺ late phase of *Bordetella* infection.

15 55. The immunogenic composition of claims 1-54 comprising a polypeptide that is expressed during the Bvgⁱ phase of *Bordetella* infection.

56. The immunogenic composition of claims 1-55 comprising an antigen that is expressed during the Bvg⁻ phase of *Bordetella* infection.

20 57. The immunogenic composition of claims 1-56, further comprising diphtheria toxoid and tetanus toxoid.

25 58. The immunogenic composition of claim 1-57 further comprising PRP capsular oligosaccharide or polysaccharide from *Haemophilus influenzae* b, preferably conjugated to a source of T-cell epitopes.

59. The immunogenic composition of claim 1-58 further comprising HbsAg.

30 60. The immunogenic composition of claim 1-59 further comprising IPV.

61. The immunogenic composition of claim 1-60 further comprising one or more of Men A, C, W or Y capsular polysaccharides or oligosaccharides, preferably conjugated to a source of T-cell epitopes.
- 5 62. The immunogenic composition of claims 1-61 further comprising a protein from *N. meningitidis* serogroup B.
63. The immunogenic composition of claim 1-62 further comprising one or more capsular polysaccharides or oligosaccharides from *S. pneumoniae*, preferably
10 conjugated to a source of T-cell epitopes.
64. The immunogenic composition of claim 1-63 further comprising killed attenuated Hepatitis A virus.
- 15 65. A vaccine comprising the immunogenic composition of claim 1-64 and an adjuvant.
66. Use of the immunogenic composition of claims 1-64 in the preparation of a medicament for use in generating an immune response against Bordetella infection
- 20 67. The use of claim 66 in the preparation of a medicament for use in the treatment or prevention of both *B. pertussis* and *B. parapertussis* infection.
68. A method for treating or preventing Bordetella infection comprising administering the vaccine of claim 65 to a host.
- 25 69. The method of claim 68 in which both *B. pertussis* and *B. parapertussis* infection is treated or prevented.

ABSTRACT OF THE DISCLOSURE

The invention provides BASB232 polypeptides and polynucleotides encoding BASB232
5 polypeptides and methods for producing such polypeptides by recombinant techniques.
Also provided are diagnostic, prophylactic and therapeutic uses. The invention further
provides immunogenic compositions comprising a plurality of antigens selected from at
least two different categories of antigen, having different functions within Bordetella.
Examples of such categories of antigen are autotransporter proteins, iron acquisition
10 proteins, lipoproteins, adhesins and toxins/invasins.

Figure 1

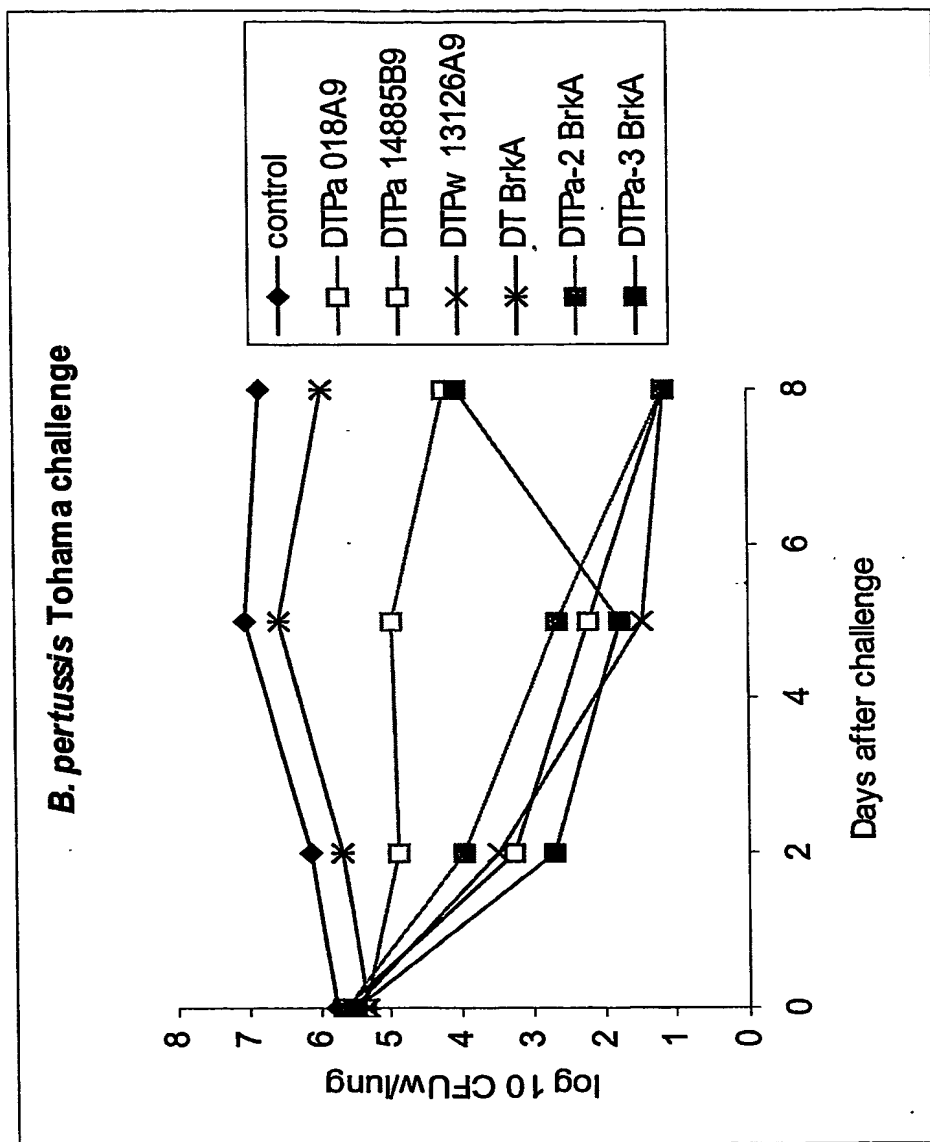


Figure 2

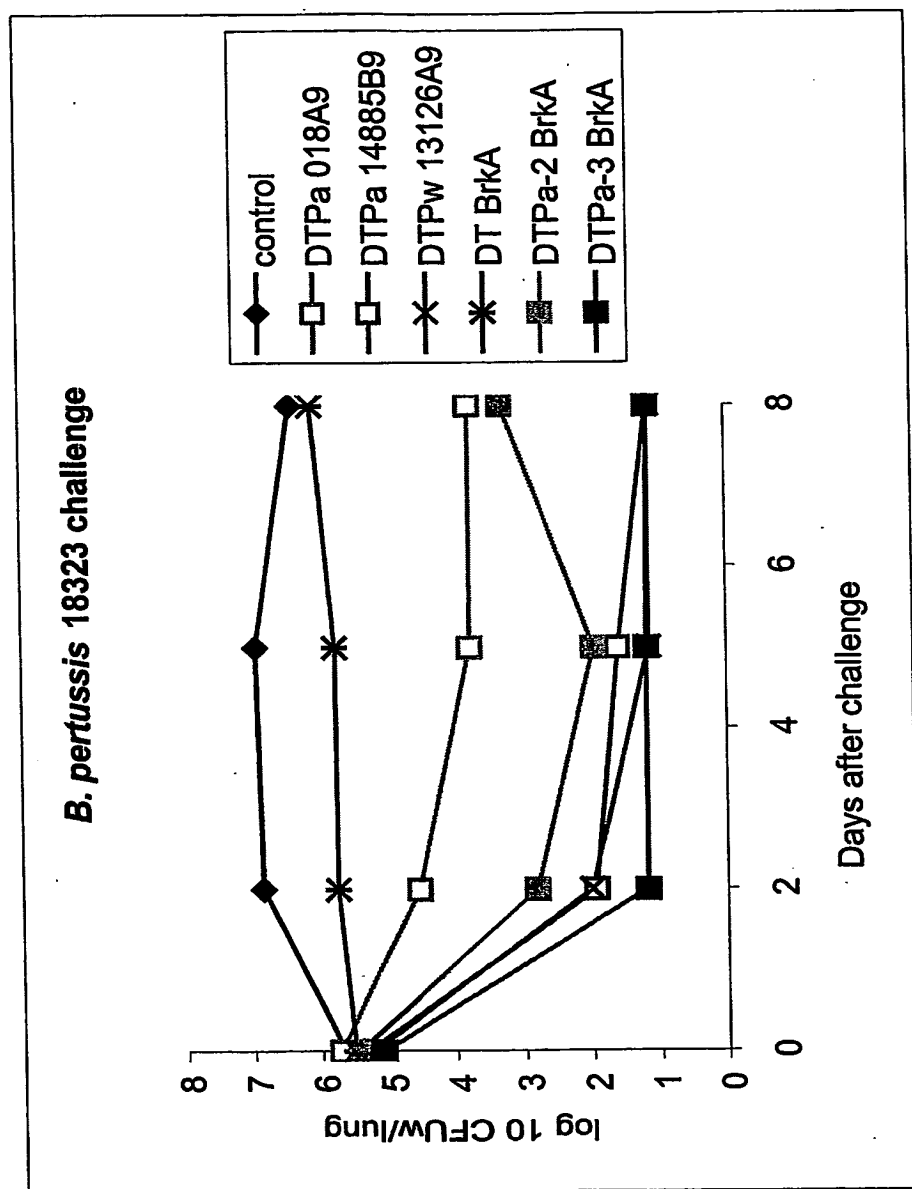
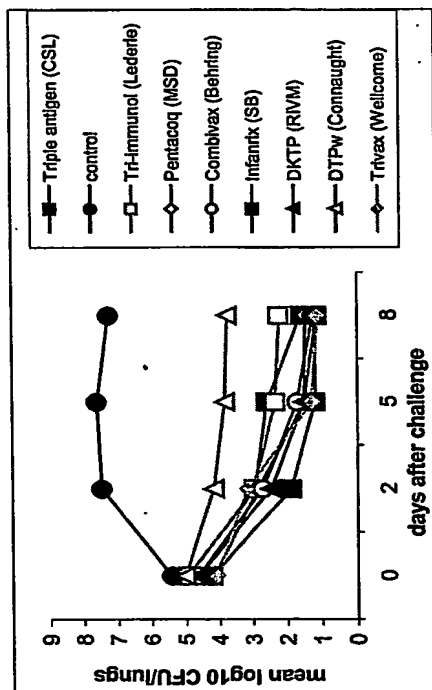


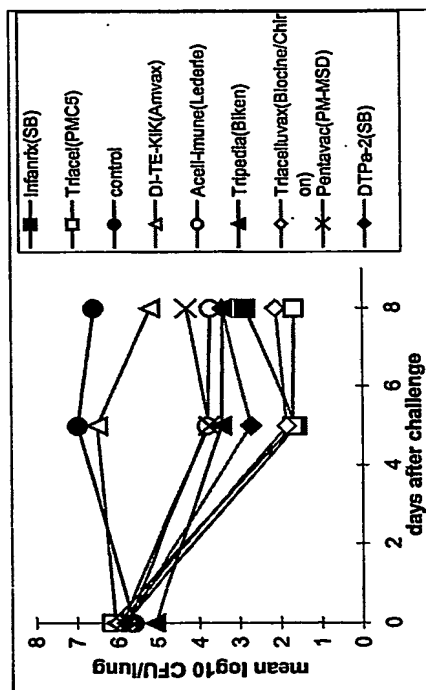
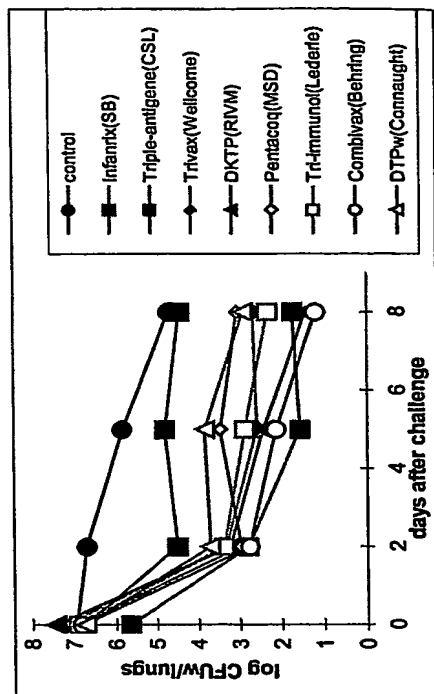
Figure 3

B. pertussis

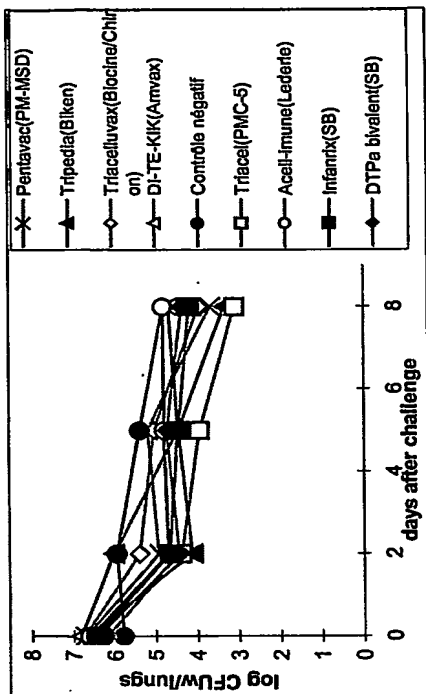


DTPw

B. parapertussis



DTPa



2
PL

PCT/EP2004/011082

